

us-09-455-486-6.Tag

OM protein - protein search, using sw model

(without alignments)
1336.214 Million cell updates/sec

Scoring table: BLOSUM62

Total number of hits satisfying chosen parameters: 1586107

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

Result No.	Score	Query Match	length	DB	ID	Description
1	2351	100.0	454	4	AAE02781	Aae02781 Human six
2	2351	100.0	454	4	AAU10188	Aau10188 Human ORF
3	2351	100.0	454	5	ABG61933	Abg61933 Prostate
4	2351	100.0	454	5	AAU76538	Aau76538 Tumour-as
5	2351	100.0	454	5	AAU80190	Aau80190 Human PUM
6	2351	100.0	454	7	ABU08893	Abu08893 Tumour-as
7	2348	99.9	454	7	ADB65001	Adb65001 Human pro
8	2294	97.6	490	5	AAE28951	Aae28951 Human STE
9	2294	97.6	490	7	ABU63312	Abu63312 Human tra
10	2290	97.4	490	7	AAU10187	Aau10187 Human six
11	2036	86.6	419	4	AAU10189	Aau10189 Human ORF
12	1972.5	83.9	576	4	ABG12306	Abg12306 Novel hum
13	1972.5	83.9	1273	4	ABG00113	Abg00113 Novel hum
14	1738	73.9	450	4	AAE02841	Aae02841 Human STE
15	1364.5	58.0	1082	5	ABP62883	Abp62883 Human pol
16	1316	56.0	488	4	AAE49483	Aad49483 Rat p-HYD
17	1316	56.0	526	5	ABH83365	Abh83365 Murine Tu
18	1272	54.1	488	4	AAH85775	Aab85775 Human diru
19	1272	54.1	488	5	ABH83366	Abh83366 Human Tum
20	1269	54.0	488	4	AAH93224	Aab93224 Human pro
21	1269	54.0	488	4	AAU04564	Aau04564 Human G-p
22	1269	54.0	488	4	AAU10220	Aau10220 Human Six
23	1269	54.0	488	6	ABU60864	Abu60864 Human G p
24	1245.5	53.0	487	4	AAH49481	Aab49481 Human p-H4
25	1088	46.3	459	4	AAH74715	Aab74715 Human mem

26	1087	46.2	459	4	AAE02782	Aae02782	Human	slx
27	1087	46.2	459	5	ABP64820	Abp64820	Human	pro
28	1085	46.2	459	4	AAU10190	Aau10190	Human	six
29	1082	46.0	458	4	AAE02636	Aae02636	Human	STE
30	1074.5	45.7	456	4	AAB49482	Aab49482	Human	p-h
31	901	38.3	173	3	AAy58195	Aay58195	Human	STR
32	844	35.9	179	4	AAm18006	Aam18006	Peptide	#
33	844	35.9	179	4	ABB37041	Abb37041	Peptide	#
34	844	35.9	179	4	AAm30517	Aam30517	Peptide	#
35	844	35.9	179	4	ABB31810	Abb31810	Peptide	#
36	844	35.9	179	4	ABB22356	Abb22356	Protein	#
37	844	35.9	179	4	AAm70183	Aam70183	Human	bon
38	844	35.9	179	4	AAm57768	Aam57768	Human	bra
39	844	35.9	179	4	ABG51883	Abg51883	Human	liv
40	844	35.9	179	4	AAm05646	Aam05646	Peptide	#
41	844	35.9	179	5	ABG39817	Abg39817	Human	pep
42	768	32.7	237	4	AAU04565	Aau04565	Human	G-p
43	768	32.7	237	6	ABU60865	Abu60865	Human	G
44	736	31.3	141	3	AAy52589	Aay52589	Human	pro
45	717	30.5	267	6	ABU60886	Abu60886	Human	G p

RESULT 1

ID	AAE02781	standard; protein; 454 AA.
XX	AAE02781;	
AC	06-AUG-2001	(first entry)
XX		
DT	Human six transmembrane epithelial antigen of pro	
DE		
XX	Human; cytosolic; antiproliferative; vaccine; ge	
KW	six transmembrane epithelial antigen of the prost	
KW	chromosome 7q21; cancer; prostate; colon; bladder	
KV	pancreatic.	
XX		
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	Region	100..108
FT		/label= HLA-A2_binding_peptide #5
FT	Region	153..165
FT		/label= Immunogenic_peptide #1
FT	Domain	207..228
FT		/label= Transmembrane_domain #1
FT	Region	227..235
FT		/label= HLA-A2_binding_peptide #1
FT	Domain	255..277
FT		/label= Transmembrane_domain #2
FT	Domain	304..325
FT		/label= Transmembrane_domain #3
FT	Region	306..314
FT		/label= HLA-A2_binding_peptide #4
FT	Region	307..315
FT		/label= HLA-A2_binding_peptide #3
FT	Region	345..358
FT		/label= Immunogenic_peptide #2
FT	Domain	359..381
FT		/label= Transmembrane_domain #4
FT	Domain	393..415
FT		/label= Transmembrane_domain #5
FT	Region	402..410
FT		/label= HLA-A2_binding_peptide #2
FT	Domain	428..450
FT		/label= Transmembrane_domain #6
XX		
PN	W0200140276-A2.	
XX		
PD	07-JUN-2001.	

XX 06-DEC-2000; 2000WO-US033040.
PF
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Afar DEH, Hubert RS, Raitano AB, Safiran DC, Mitchell SC;
PI Faris M, Jakobovits A;
XX
XX WPI; 2001-367804/38.
DR N-PSDB; AAD07072.
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
XX Claim 1; Fig 9A-9C; 187pp; English.
XX
CC The present sequence is human six transmembrane epithelial antigen of the
CC prostate (STEAP)-2 protein. STEAP is a member of cell surface serpentine
CC transmembrane antigens. STEAP-2 gene is located on chromosome 7q21 and is
CC used in gene therapy. Inhibiting the development or progression of a
CC cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic)
CC expressing STEAP or inhibiting growth or killing cells expressing STEAP
CC in a patient, comprises administering a vaccine composition to the
CC patient. Treating a patient with a cancer that expresses STEAP, or
CC inhibiting growth or killing cells expressing STEAP, comprises
CC administering to the patient a vector encoding single chain monoclonal
CC antibody that comprises the variable domains of the heavy and light
CC chains of the monoclonal antibody that specifically binds to STEAP, such
CC that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02841) in sequence
CC listing of the specification. However both the sequences differ at
CC several positions
XX
SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMNGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
DB 1 MESISMNGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
QY 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120
QY 121 RINGYBESNAEYLASLFPDSLIVKGFNVSAWALQGPKASROVYICSNNOARQVIE 180
DB 121 RINGYBESNAEYLASLFPDSLIVKGFNVSAWALQGPKASROVYICSNNOARQVIE 180
QY 181 LARQLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFYRDVTHPYA 240
DB 181 LARQLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFYRDVTHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATLLSLVYIAGLLAAAYQLYGTXYRRPFWLETWLO 300
DB 241 RNOQSDFYKPIEIVNKTLPVATLLSLVYIAGLLAAAYQLYGTXYRRPFWLETWLO 300
QY 301 CRKOLGLISFFPAMVAVASLCPLMRSEERYLFLNMAYQOVHANIEENSWNEEYVRIEM 360
DB 301 CRKOLGLISFFPAMVAVASLCPLMRSEERYLFLNMAYQOVHANIEENSWNEEYVRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTYTPNFVLAIVLPSIVIDLQLCRYPD 454

DB 421 EBYRYFTYTPNFVLAIVLPSIVIDLQLCRYPD 454
RESULT 2
ID AAU10188 standard; protein; 454 AA.
XX
XX AC AAU10188;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human ORF2 of Six-Transmembrane Protein of Prostate 1.
XX
KW Human; Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF2.
XX
OS Homo sapiens.
XX
PN WO200172962-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009410.
XX
PR 24-MAR-2000; 2000US-0191929P.
XX
PA (SAAT/) SAATCIOGLU F.
XX
PI Saatioglu F;
XX
DR WPI; 2001-662926/76.
DR N-PSDB; AAS15810, AAS15811.
XX
PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.
XX
PS Claim 1; Fig 4H; 114pp; English.
XX
XX The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence is prostate specific protein, Six-Transmembrane
CC Protein of Prostate 1, STMP1, ORF2
XX
SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMNGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
DB 1 MESISMNGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
QY 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120

QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIOARQOYIE 180
DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIOARQOYIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLISLVYLAGLLAAAYQLYGTYKRRFPPLLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLISLVYLAGLLAAAYQLYGTYKRRFPPLLETWLQ 300
QY 301 CRKQGLISFPFAMVHVAYSCLPMRSEERYLFLNMAYYQOVHANIENSWNEEYVRIEMY 360
DB 301 CRKQGLISFPFAMVHVAYSCLPMRSEERYLFLNMAYYQOVHANIENSWNEEYVRIEMY 360
QY 361 ISFGIMSLGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 BEYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454
DB 421 BEYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454

RESULT 3
ABG61933

ID ABG61933 standard; protein; 454 AA.

XX ABG61933;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated protein #134.

KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

OS Mammalia.

PN W0200230268-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US032045.

PR 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.

PR 16-MAR-2001; 2001US-0276791P.

PR 16-MAR-2001; 2001US-0276888P.

PR 06-APR-2001; 2001US-0281922P.

PR 24-APR-2001; 2001US-0286214P.

PR 30-APR-2001; 2001US-00847046.

PR 04-MAY-2001; 2001US-0288589P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

DR WPI; 2002-471335/50.

DR N-PSDB; ABK92252.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.

PS Claim 27; Page 416; 436pp; English.

CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSISETCLPNGINGIKARKVTVGIVIGSGDFAKSLTIRLIRCGYHVIGS 60
DB 1 MESISMGSPKSISETCLPNGINGIKARKVTVGIVIGSGDFAKSLTIRLIRCGYHVIGS 60
QY 61 RNPKFASEFPFHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
DB 61 RNPKFASEFPFHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIOARQOYIE 180
DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIOARQOYIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLISLVYLAGLLAAAYQLYGTYKRRFPPLLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLISLVYLAGLLAAAYQLYGTYKRRFPPLLETWLQ 300
QY 301 CRKQGLISFPFAMVHVAYSCLPMRSEERYLFLNMAYYQOVHANIENSWNEEYVRIEMY 360
DB 301 CRKQGLISFPFAMVHVAYSCLPMRSEERYLFLNMAYYQOVHANIENSWNEEYVRIEMY 360
QY 361 ISFGIMSLGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 BEYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454
DB 421 BEYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454

RESULT 4
AAU76538

ID AAU76538 standard; protein; 454 AA.

AC AAU76538;

DT 05-JUN-2002 (first entry)

DE Tumour-associated antigenic target protein, TAT138.

KW TAT138; Tumour-associated Antigenic Target; tumour; breast cancer;

KW colorectal cancer; lung cancer; ovarian cancer;

KW central nervous system cancer; liver cancer; bladder cancer; melanoma;

XX pancreatic cancer; leukaemia; gene therapy.

OS Homo sapiens.

FT Modified-site 8.13

FT Modified-site /note= "N-myristoylation site"

FT Modified-site 24.29

FT /note= "N-myristoylation site"
FT 29..32
FT /note= "cAMP- and cGMP-dependent protein kinase
FT phosphorylation site"
FT 34..39
FT /note= "N-myristoylation site"
FT 193..198
FT /note= "N-myristoylation site"
FT 210..230
FT /note= "Transmembrane domain"
FT 256..259
FT /note= "Asn is N-glycosylated"
FT 257..277
FT /note= "Transmembrane domain"
FT 259..379
FT /note= "Transmembrane domain"
FT 274..279
FT /note= "N-myristoylation site"
FT 299..319
FT /note= "Transmembrane domain"
FT 393..413
FT /note= "Transmembrane domain"
FT 416..424
FT /note= "Tyrosine kinase phosphorylation site"
FT 428..448
FT /note= "Transmembrane domain"
XX WO200216429-A2.
XX 28-FEB-2002.
XX 22-JUN-2001; 2001WO-US020118.
XX 24-AUG-2000; 2000WO-US023328.
XX 26-SEP-2000; 2000US-0235451P.
XX 01-DEC-2000; 2000WO-US032678.
XX 28-FEB-2001; 2001WO-US006520.
XX 01-MAR-2001; 2001WO-US006666.
XX (GETH) GENENTECH INC.
XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
XX Wood WI, Wu TD, Zhang Z;
XX WPI; 2002-280917/32.
XX N-PSDB; ABK11093.
XX Novel isolated tumor-associated antigenic target polypeptides which are
XX useful as targets for cancer therapy and diagnosis in mammals.
XX Claim 12; Fig 10; 121pp; English.
XX The invention relates to an isolated tumour-associated antigenic target
XX polypeptide (TAT) (I), specifically TAT134-TAT138 polypeptides, and the
XX polynucleotides (II) encoding them. (II) is useful for diagnosing the
XX presence of a tumour in a mammal, where the level of expression of (II)
XX is indicative on the presence of tumour in the mammal from which the test
XX sample was obtained. Antibody to (I) is useful for killing a cancer cell,
XX (e.g. breast cancer cell, a colorectal cancer cell, a lung cancer cell,
XX an ovarian cancer cell, a central nervous system (CNS) cancer cell, a
XX liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a
XX melanoma cell or a leukaemia cell) that expresses (I). Oligonucleotides
XX hybridising to (II) are useful as diagnostic probes, antisense
XX oligonucleotide probes or for encoding fragments of full length TAT
XX polypeptide. (II) is also useful in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA probes, for constructing and in
XX hybridisation probes for mapping the gene encoding TAT and for genetic
XX analysis of individuals with genetic disorders. (II) is also useful for
XX generating either transgenic animals or knockout animals, and in gene
XX therapy. The TAT polypeptides and nucleic acids may also be used for
XX tissue typing and the TAT polypeptides are useful for screening compounds
XX that mimic the TAT polypeptide (agonist) or prevent the effect of TAT
XX polypeptide (antagonist). The antibody is useful for staging TAT

CC polypeptide-expressing cancers, purifying or immunoprecipitating TAT
CC polypeptide from cells, for detection and quantitation of TAT polypeptide
CC in vitro, e.g., in an enzyme linked immunosorbent assay (ELISA) or
CC Western blot. The antibodies are also useful for treating a TAT-
CC expressing cancer or alleviating one or more symptoms of cancer in a
CC mammal. The present sequence represents the amino acid sequence of TAT138
XX
SQ Sequence 454 AA;
Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. NO. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESISMWGPSKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHWIGS 60
DB 1 MESISMWGPSKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHWIGS 60
QY 61 RNPKEASEFPFHVVDVTHHEDALTKTNIIFFVAIHREHYTSLMDLRHLVGGKLLIDVSNM 120
DB 61 RNPKEASEFPFHVVDVTHHEDALTKTNIIFFVAIHREHYTSLMDLRHLVGGKLLIDVSNM 120
QY 121 RINQYBESNAEYLASLFPDILLVKGFNVSAMALQLGPKDASROVYICSNNOARQVIE 180
DB 121 RINQYBESNAEYLASLFPDILLVKGFNVSAMALQLGPKDASROVYICSNNOARQVIE 180
QY 181 LARQLNFIPIDLGSLSSAREIENLPRLFTLMRGVYVAISLATFFFLYSFVDVIHPYA 240
DB 181 LARQLNFIPIDLGSLSSAREIENLPRLFTLMRGVYVAISLATFFFLYSFVDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
QY 301 CRKQIGLLSFPFAMVHAYSLCLPMRBSERYLFNNAYQOVHANIENTSNNEEYWRILEMY 360
DB 301 CRKQIGLLSFPFAMVHAYSLCLPMRBSERYLFNNAYQOVHANIENTSNNEEYWRILEMY 360
QY 361 ISFGIMSLGLSLLAIVTSIPSVSNALNWRBPSFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLSLLAIVTSIPSVSNALNWRBPSFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EEYRFFYTPPNFVLAIVPSIVILDLQLCRYPD 454
DB 421 EEYRFFYTPPNFVLAIVPSIVILDLQLCRYPD 454
RESULT 5
AAU80190
ID AAU80190 standard; protein; 454 AA.
XX
XX AAU80190;
AC
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PUMPCn protein, PRO23203.
KW Human; protein upregulated in metastatic prostate cancer; immunogen;
KW PUMPCn; PRO23203; prostate cancer; cytostatic; gene therapy; transgenic;
KW androgen independent prostate cancer; DNA microarray.
OS Homo sapiens.
XX
XX WO200226822-A2.
PN
XX
PD 04-APR-2002.
XX
PF 26-SEP-2001; 2001WO-US030290.
XX
PR 26-SEP-2000; 2000US-0235451P.
XX
PA (GETH) GENENTECH INC.
XX
PI Devaux B, Eberhard D, Goddard A, Godowski PJ, Grimaldi JC;

PI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
XX WPI; 2002-383270/41.
DR N-PSDB; ABK50391.

XX New polypeptide termed protein upregulated in metastatic prostate cancer
PT and encoding polynucleotides, useful for identifying polypeptide
PT antagonists for treating prostate cancer.

PS Claim 23; Fig 2; 137pp; English.

XX The invention relates to an isolated human protein upregulated in
CC metastatic prostate cancer (PUMPCn) PRO23203 polypeptide, a sequence 80%
CC identical to PRO23203 and the sequence as encoded by cDNA insert of the
CC vector deposited as ATCC Deposit No. PTA-2513 (DNA185171-2994) on
CC 26/9/2000. Also included are the polynucleotide encoding the protein (or
CC a DNA sequence 80% identical to the polynucleotide and one that
CC hybridises to complement of the polynucleotide) and one that
CC polynucleotide, a polynucleotide deposited with ATCC under accession
CC number PTA-2513 (DNA185171-2994), a host cell comprising the vector,
CC preparation of PRO23203, a chimeric molecule comprising PRO23203 fused to
CC a heterologous amino acid sequence, an anti-PRO23203 antibody, an
CC agonist/antagonist of PRO23203, and diagnosing the presence of prostate
CC cancer in a mammal by: (a) contacting a microarray diagnostic with a
CC DNA185171-2994 probe, detecting and quantifying hybridisation of
CC tissue and determining if DNA185171-2994 is overexpressed; or (b)
CC contacting a tissue of the mammal with an anti-PRO23203 antibody and
CC detecting the binding of the antibody to a component of the tissue, where
CC binding is indicative of the presence of prostate cancer in the mammal.
CC The antibody is useful for treating prostate cancer in mammal which is
CC androgen independent prostate cancer, that has metastasised to another
CC portion of the body, where the antibody is not conjugated with a
CC cytotoxic agent and the method further comprises administering a
CC chemotherapeutic agent to the mammal and for diagnosing the presence of
CC prostate cancer in a mammal. The PRO23203 polynucleotide is useful for
CC generating either transgenic animals or knock out animals which in turn
CC are useful in the development and screening of therapeutically useful
CC reagents or for use in gene therapy and for chromosome identification.
CC PRO23203 and polynucleotide may also be used for tissue typing and
CC PRO23203 may also be employed as a therapeutic agents and for screening
CC compounds to identify those that mimic the PRO23203 polypeptides
CC (agonists) or prevent the effect of the PRO23203 polypeptide
CC (antagonists). PRO23203 is further useful for the affinity purification
CC of PRO23203 from recombinant cell culture or natural sources. The present
CC sequence represents PRO23203

XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMMSGPKSLSETCLPNGINGIKAKKVTGVISSGDFAKSLTIRLCGYHVTGS 60
DB 1 MESISMMSGPKSLSETCLPNGINGIKAKKVTGVISSGDFAKSLTIRLCGYHVTGS 60
QY 61 RNPKFASSEFPFHVVDVTHHEDALTKNTIIFVAIHRHHTSLMDLRHLLVSKILIDVSNM 120
DB 61 RNPKFASSEFPFHVVDVTHHEDALTKNTIIFVAIHRHHTSLMDLRHLLVSKILIDVSNM 120
QY 121 RINQYPPESNAEYIASLFPDSLIVKGFNVSAWALQGPDKASROYICSNMIQARQVTE 180
DB 121 RINQYPPESNAEYIASLFPDSLIVKGFNVSAWALQGPDKASROYICSNMIQARQVTE 180
QY 161 LARQINFIPIIDGSLSSAREIENPLRLFTLWKGPPVVAISLATFFFLYSFVRDVIHPYA 240
DB 161 LARQINFIPIIDGSLSSAREIENPLRLFTLWKGPPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPILIVNKTLPVIAITLLSLVLAGLLAAAYQLYGTGKRRFPWLEIWLQ 300
DB 241 RNQOSDFYKIPILIVNKTLPVIAITLLSLVLAGLLAAAYQLYGTGKRRFPWLEIWLQ 300

QY 301 CRKQGLLSFPFAMVHVAYSLCLPMRRSERVFLNMAVQOVHANINSEWNEEVRRIEM 360
DB 301 CRKQGLLSFPFAMVHVAYSLCLPMRRSERVFLNMAVQOVHANINSEWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLIAVTSIPSVSNALNWEESFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLSLIAVTSIPSVSNALNWEESFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EBYRFTYPPNFVIALVPSIVILDLQLCRYPD 454
DB 421 EBYRFTYPPNFVIALVPSIVILDLQLCRYPD 454

RESULT 6
ABU08893
ID ABU08893 standard; protein; 454 AA.
XX
AC ABU08893;
DT 09-OCT-2003 (first entry)
XX
DE Tumour-associated antigenic target 138 (TAT138).

XX Human; tumour-associated antigenic target polypeptide; TAT; tumour;
XX gene therapy; cytostatic; tissue typing; prostate tumour; cancer.
XX Homo sapiens.

US2003060612-A1.

27-MAR-2003.

22-JUN-2001; 2001US-00888257.

PR 28-OCT-1997; 97US-0063540P.
PR 17-JUN-1998; 98US-0089653P.
PR 10-SEP-1998; 98US-0099792P.
PR 08-OCT-1998; 98US-0103678P.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 18-FEB-2000; 2000WO-US004342.
PR 02-MAR-2000; 2000WO-US005841.
PR 30-MAR-2000; 2000WO-US008439.
PR 24-AUG-2000; 2000WO-US023328.
PR 26-SEP-2000; 2000US-0235451P.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.

PA (GETH) GENENTECH INC.

PI Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
PI Wood WI, Wu TD, Zhang Z;

WPI; 2003-576355/54.
DR N-PSDB; ACD25893.

PT New nucleic acid, useful for the manufacture of a medicament for
PT diagnosing or treating tumor in a mammal.

PS Claim 12; Fig 10; 71pp; English.

XX The invention discloses human nucleic acids encoding tumour-associated
CC antigenic target (TAT) polypeptides, with or without their associated
CC signal peptide. Also disclosed is an antibody that specifically binds to
CC the TAT polypeptides, a method for detecting the presence of a tumour in
CC a mammal and a method for killing a cancer cell expressing the TAT
CC polypeptide. The nucleotide sequences are useful in preparing TAT
CC replacement of defective gene). The TAT polypeptides are useful as
CC therapeutic agents and for detecting the presence, prevention and/or
CC treatment of a tumour, such as colon, breast or prostate tumour. The TAT
CC polypeptides and nucleic acids may also be used diagnostically for tissue

No prior art

CC typing. The sequence presented is the TAT138 polypeptide of the invention
XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
DB 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
QY 61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120
QY 121 RINQYPESNAEYTLASLPDLSLVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYTLASLPDLSLVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLLETWLQ 300
QY 301 CRKQLGLSFFPAMVHVAAYSLCLPMRSEERYFLNMAYQOVHANIENTSNWNEEVRRIEM 360
DB 301 CRKQLGLSFFPAMVHVAAYSLCLPMRSEERYFLNMAYQOVHANIENTSNWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHYLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHYLIYGWKRAFE 420
QY 421 EBYRFTYTPNFTLVLPISIVILDLLQLCRYPD 454
DB 421 EBYRFTYTPNFTLVLPISIVILDLLQLCRYPD 454

RESULT 7
ADB65001
ID ADB65001 standard; protein; 454 AA.

XX ADB65001;
DT 04-DEC-2003 (first entry)

DE Human protein encoded by clone PROST20168600.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.

OS Homo sapiens.

PN EP1308459-A2.

PD 07-MAY-2003.

PF 28-MAR-2002; 2002EP-00007401.

PR 05-NOV-2001; 2001JP-00379298.

PR 25-JAN-2002; 2002US-00350978.

PA (HELI-) HELIX RES INST.
(REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;

XX WPI; 2003-450961/43.
DR N-PSDB; ADB63031.
XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

PS The invention discloses a polynucleotide comprising a sequence selected
XX from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.

XX Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.1e-244;
Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
DB 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60

QY 61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120

QY 121 RINQYPESNAEYTLASLPDLSLVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYTLASLPDLSLVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180

QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLLETWLQ 300

QY 301 CRKQLGLSFFPAMVHVAAYSLCLPMRSEERYFLNMAYQOVHANIENTSNWNEEVRRIEM 360
DB 301 CRKQLGLSFFPAMVHVAAYSLCLPMRSEERYFLNMAYQOVHANIENTSNWNEEVRRIEM 360

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHYLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHYLIYGWKRAFE 420

QY 421 EBYRFTYTPNFTLVLPISIVILDLLQLCRYPD 454
DB 421 EBYRFTYTPNFTLVLPISIVILDLLQLCRYPD 454

RESULT 8
AAE28951
ID AAE28951 standard; protein; 490 AA.
XX AC AAE28951;
XX DT 27-JAN-2003 (first entry)
XX DE Human STEAPRP protein.
XX KW Human; STEAPRP; proliferative disorder; prostate hyperplasia; cancer;
KM six-transmembrane epithelial antigen of the prostate related protein;
KW gene therapy; cytostatic.
XX OS Homo sapiens.
XX PN WO200272596-A1.
XX PD 19-SEP-2002.
XX PF 07-MAR-2002; 2002WO-US007053.
XX PR 09-MAR-2001; 2001US-00802520.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Lal PG, Faris M, Chen H, Ison CH;
XX DR WPI; 2002-723311/78.
XX DR N-PSDB; AAD46429.
XX PT New six-transmembrane epithelial antigen of the prostate-related protein
PT and nucleic acids, useful for diagnosing and treating prostate cell
PT proliferative disorders, particularly prostate hyperplasia and prostate
PT cancer.
XX PS Claim 13; Page 54-55; 62pp; English.
XX CC The invention relates to six-transmembrane epithelial antigen of the
CC prostate related protein (STEAPRP) and its corresponding nucleic acid.
CC STEAPRP DNA is used to diagnose and to treat prostate cell proliferative
CC disorders, particularly prostate hyperplasia and prostate cancer, and for
CC screening a library of molecules of compounds for specific binding
CC affinity. It is also used in gene therapy. STEAPRP is used to screen a
CC subject sample for antibodies, which specifically binds the protein, and
CC to prepare and purify a protein. The present sequence is human STEAPRP
CC protein
XX SQ Sequence 490 AA;
Query Match 97.6%; Score 2294; DB 5; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MESISMNGSPKSLSETCLPENGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVIGS 60
DB 1 MESISMNGSPKSLSETCLPENGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVIGS 60
OY 61 RNPKRASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
DB 61 RNPKRASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
OY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASRQVYICSNICARQOVIE 180
DB 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASRQVYICSNICARQOVIE 180
OY 181 LARQLNFIPIDLGLSSAREIENIPRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
DB 181 LARQLNFIPIDLGLSSAREIENIPRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
OY 241 RNQGSDFYKIPRIEIVNKTLPVATLLSLVYIAGLLAAAYQLYGTGKRRFPPELWTWQ 300
DB 241 RNQGSDFYKIPRIEIVNKTLPVATLLSLVYIAGLLAAAYQLYGTGKRRFPPELWTWQ 300

DB 241 RNQGSDFYKIPRIEIVNKTLPVATLLSLVYIAGLLAAAYQLYGTGKRRFPPELWTWQ 300
OY 301 CRKQLGLISFFPAMVHVAYSLCPMRSERYLFLNMAYQVHANIENSWNEEVRIMY 360
DB 301 CRKQLGLISFFPAMVHVAYSLCPMRSERYLFLNMAYQVHANIENSWNEEVRIMY 360
OY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
OY 421 EBYRFRYTPPNFVLALVLPISVIL 444
DB 421 EBYRFRYTPPNFVLALVLPISVIL 444
RESULT 9
ABU63312
ID ABU63312 standard; protein; 490 AA.
XX AC ABU63312;
XX DT 18-SEP-2003 (first entry)
XX DE Human transmembrane protein TMPL.
XX KW Human; prostate tumour; lung tumour; TMPL; cancer.
XX OS Homo sapiens.
XX PN US2003064397-A1.
XX PD 03-APR-2003.
XX PF 24-JUL-2002; 2002US-00205267.
XX PR 22-MAY-1998; 98US-00083521.
XX PR 16-SEP-1999; 99US-00397558.
XX PR 09-MAR-2001; 2001US-00802520.
XX PR 26-SEP-2001; 2001US-00963896.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Spancake KM, Rickert PK, Lal PG, Ison CH;
XX DR WPI; 2003-540783/51.
XX DR N-PSDB; ACD27642.
XX PT New transmembrane protein differentially expressed in prostate and lung
PT tumors, useful for diagnosing, staging, treating or monitoring
PT progression or treatment of cancer, preferably lung or prostate cancer.
XX PS Claim 1; Fig 1; 49pp; English.
XX CC The invention relates to a purified human protein comprising a
XX polypeptide having a fully defined transmembrane protein differentially
XX expressed in prostate and lung tumour (TMPL) sequence. The protein is
XX useful for detecting expression of the protein in a sample, where the
XX sample is lung or prostate tissue sample and the protein is
XX differentially expressed when compared with a standard and is diagnostic
XX of a lung or prostate cancer. The protein is useful to identify an
XX antibody that specifically binds the protein. The protein is useful for
XX screening several molecules and compounds to identify at least one
XX ligand. The protein is also useful for preparing and purifying a
XX polyclonal antibody and for preparing a monoclonal antibody using
XX standard animal immunisation/hybridoma methods. A TMPL specific antibody
XX is useful to detect expression of the protein in a lung or prostate
XX tissue sample and the method is useful as a diagnostic for lung or
XX prostate cancer. A TMPL specific antibody is also useful for
XX immunopurifying a protein, for treating a prostate cancer and for
XX delivering a therapeutic agent to a cancer, preferably prostate cancer.
XX The protein is useful for diagnosing, staging, treating or monitoring
XX progression or treatment of cancer, preferably lung or prostate cancer.
XX The present sequence represents the amino acid sequence of the human

CC transmembrane protein differentially expressed in prostate and lung
CC tumour TMPL
XX
SQ Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSECTLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60
Db 1 MESISMGSPKSLSECTLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60
QY 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAVALQGPXDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAVALQGPXDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQNFIPIDLSLSAREIENLPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKLPIVAITLLSLVYLAGLLAAAYQLYGTRRRPFWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKLPIVAITLLSLVYLAGLLAAAYQLYGTRRRPFWLETWLQ 300
QY 301 CRKQLGLSFFFAVHVAVSLCLPMRSERYFLNMAVYQOVHANINSNWNEEVRRIEM 360
Db 301 CRKQLGLSFFFAVHVAVSLCLPMRSERYFLNMAVYQOVHANINSNWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EBYRFFYTPPNFVLALVPSIVIL 444
Db 421 EBYRFFYTPPNFVLALVPSIVIL 444

RESULT 10
AAU10187
ID AAU10187 standard; protein; 490 AA.
AC AAU10187;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human Six-Transmembrane Protein of Prostate 1, STMPL.
XX
KW Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH Domain 209..230
FT /label= Transmembrane_domain_1
FT Domain 255..273
FT /label= Transmembrane_domain_2
FT Domain 304..325
FT /label= Transmembrane_domain_3
FT Domain 360..380
FT /label= Transmembrane_domain_4
FT Domain 393..415
FT /label= Transmembrane_domain_5
FT Domain 432..452
FT /label= Transmembrane_domain_6

XX
PN WO200172962-A2.
XX
PD 04-OCT-2001.
XX
PE 23-MAR-2001; 2001WO-US009410.
XX
PR 24-MAR-2000; 2000US-0191929P.
XX
PA (SAAT/) SAATCIOGLU F.
PI SaaticioGLU F;
XX

DR WPI; 2001-662926/76.
DR N-PSDB; AAS15793, AAS15801, AAS15802.
XX
PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.

PS Claim 1; Fig 4B; 114pp; English.

CC The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence represents a prostate specific protein, Six-
CC Transmembrane Protein of Prostate 1, STM;1

XX
SQ Sequence 490 AA;

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 2.3e-238;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSECTLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60
Db 1 MESISMGSPKSLSECTLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60
QY 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAVALQGPXDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAVALQGPXDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQNFIPIDLSLSAREIENLPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKLPIVAITLLSLVYLAGLLAAAYQLYGTRRRPFWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKLPIVAITLLSLVYLAGLLAAAYQLYGTRRRPFWLETWLQ 300
QY 301 CRKQLGLSFFFAVHVAVSLCLPMRSERYFLNMAVYQOVHANINSNWNEEVRRIEM 360
Db 301 CRKQLGLSFFFAVHVAVSLCLPMRSERYFLNMAVYQOVHANINSNWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EBYRFFYTPPNFVLALVPSIVIL 444

DB 421 EEYRYFTPENFVLALVLPISIVL 444

RESULT 11

AAU10189

ID AAU10189 standard; protein; 419 AA.

XX

AC AAU10189;

XX

DT 16-JAN-2002 (first entry)

XX

DE Human ORF3 of Six-Transmembrane Protein of Prostate 1.

XX

KW Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer; benign prostatic hyperplasia; acute prostatitis; testicular cancer; Kryptorchidism; testicular disorder; proliferative disorder; lymphoma; leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer; liver cancer; lung cancer; cytostatic; ORF3.

XX

OS Homo sapiens.

XX

PN WO200172962-A2.

XX

PD 04-OCT-2001.

XX

PF 23-MAR-2001; 2001WO-US009410.

XX

PR 24-MAR-2000; 2000US-0191929P.

XX

PA (SAAT/) SAATCIOGLU F.

XX

PI Saatioglu F;

XX

DR MPI; 2001-662926/76.

DR N-PSDB; AAS15813.

XX

PT New polynucleotide for the diagnosis, prevention and treatment for prostate and testis disorders, particularly prostate cancer, comprises prostate-specific or testis-specific nucleic acids.

PT

XX

PS Claim 1; Fig 4K; 114pp; English.

XX

CC The invention relates to substantially pure prostate-specific or testis-specific polypeptides and the nucleic acids encoding them. Also included are vectors and host cells expressing the proteins, a transgenic animal expressing the protein, antibodies against the proteins, probes for detecting the nucleic acids, antisense molecules for the nucleic acids and methods of isolating modulators of the proteins. Compounds that modulate the prostate specific or testis specific polypeptide are useful to diagnose, prevent or treat disorders of the testis or prostate particularly prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, kryptorchidism, undescended, retractile, ascending or vanished testis. Other proliferative disorders for which the modulators may be used include lymphoma, leukaemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.

CC

CC The present sequence is prostate specific protein, Six-Transmembrane Protein of Prostate 1, STM;1, ORF3

CC

XX

Seq Sequence 419 AA;

Query Match 86.6%; Score 2036; DB 4; Length 419; Best Local Similarity 99.7%; Pred. No. 5.5e-211; Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMWGPSKSLSETCLPNGINGIKDARKVTVGIGSGDFAKSLTIRLRGCHVAVIGS 60

DB 1 MESISMWGPSKSLSETCLPNGINGIKDARKVTVGIGSGDFAKSLTIRLRGCHVAVIGS 60

QY 61 RNPKFASFFPHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLVGLKILIDVSNM 120

DB 61 RNPKFASFFPHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLVGLKILIDVSNM 120

QY 121 RINQYPESNAEYLASLPDLSLIVKENVVSAMALQLGPKDASROYTICSNNIQARQOVIE 180

DB 121 RINQYPESNAEYLASLPDLSLIVKENVVSAMALQLGPKDASROYTICSNNIQARQOVIE 180

QY 181 LARQLNFIPIIDGLSSAREIENLEPLFTLMRGPVVAISLATPEFLYSFVRDVHPYA 240

DB 181 LARQLNFIPIIDGLSSAREIENLEPLFTLMRGPVVAISLATPEFLYSFVRDVHPYA 240

QY 241 RNOQSDFYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300

DB 241 RNOQSDFYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300

QY 301 CRKQLGLSFFPAMVHVAYSLCEPMRSERYLFLNMAVQOVHANIENSWNEEVMRIEMY 360

DB 301 CRKQLGLSFFPAMVHVAYSLCEPMRSERYLFLNMAVQOVHANIENSWNEEVMRIEMY 360

QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRPFSFIQ 395

DB 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRPFSFIQ 395

RESULT 12

ABG12306

ID ABG12306 standard; protein; 576 AA.

XX

AC ABG12306;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #12297.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Dmanac RT, Liu C, Tang YT;

XX

DR MPI; 2001-639362/73.

DR N-PSDB; AAS76493.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PT

XX

PS Claim 20; SEQ ID NO 42665; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 576 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 576;
Best Local Similarity 87.4%; Pred. No. 6.7e-204;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCHVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCHVIGS 60
QY 61 RNPKFASEFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPKFASEFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPPIVAITLISLVYAGLLAAAYQLYGTKYRRFPWLETWLO 300
DB 241 RNOQSDFYKPIEIVNKTLPPIVAITLISLVYAGLLAAAYQLYGTKYRRFPWLETWLO 300
QY 301 CRKQGLLSFFAMVHVAYSLCLPMRSEERYFLNMAYQOVHANIENSWNEEVRRIEMV 360
DB 301 CRKQGLLSFFAMVHVAYSLCLPMRSEERYFLNMAYQOVHANIENSWNEEVRRIEMV 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EYYRFTYTPNPFVLAIVLPISIVIDL 446
DB 421 EYYRFTYTPNPFVLAIVLPISIVIDL 446
QY 463 EYYRFTYTPNPFVLAIVLPISIVIDL 463

RESULT 13

ABG00113 ID ABG00113 standard; protein; 1273 AA.

XX AC ABG00113;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #104.

XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS64300.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 30472; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1273 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 1273;
Best Local Similarity 87.4%; Pred. No. 2.3e-203;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCHVIGS 60
DB 73 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCHVIGS 132
QY 61 RNPKFASEFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 133 RNPKFASEFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 192
QY 121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
DB 193 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 252
QY 181 LARQNFIPIDIGSLSSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 253 LARQNFIPIDIGSLSSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 312
QY 241 RNOQSDFYKPIEIVNKTLPPIVAITLISLVYAGLLAAAYQLYGTKYRRFPWLETWLO 300
DB 313 RNOQSDFYKPIEIVNKTLPPIVAITLISLVYAGLLAAAYQLYGTKYRRFPWLETWLO 372
QY 301 CRKQGLLSFFAMVHVAYSLCLPMRSEERYFLNMAYQOVHANIENSWNEEVRRIEMV 360
DB 373 CRKQGLLSFFAMVHVAYSLCLPMRSEERYFLNMAYQOVHANIENSWNEEVRRIEMV 411
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 412 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGWKRAFE 437
QY 421 EYYRFTYTPNPFVLAIVLPISIVIDL 446
DB 438 EYYRFTYTPNPFVLAIVLPISIVIDL 463

RESULT 14
AAE02841
ID AAE02841 standard; protein; 450 AA.
XX
AC AAE02841;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human STEAP-2 protein, alternative version.
XX
KW Human; cytosolic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KW pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 335..336
FT /note= "Encoded by GACTGAGCT"
FT Misc-difference 375..376
FT /note= "Encoded by AGATGAAGT"
FT Misc-difference 415..416
FT /note= "Encoded by GAGTAAGC"
FT Misc-difference 445..446
FT /note= "Encoded by ACATGAAGT"
FT Misc-difference 448..449
FT /note= "Encoded by AATTATTC"
XX
PN WO200140276-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033040.
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;
XX
DR WPI; 2001-367804/38.
DR N-PSDB; AAD07072.
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Claim 1; Page 175-176; 187pp; English.
XX
CC The present sequence is an alternative version of human six transmembrane
CC epithelial antigen of the prostate (STEAP)-2 protein. STEAP is a member
CC of cell surface serpentine transmembrane antigens. STEAP-2 gene is
CC located on chromosome 7q21 and is used in gene therapy. Inhibiting the
CC development or progression of a cancer (eg. prostate, colon, bladder,
CC lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or
CC killing cells expressing STEAP in a patient, comprises administering a
CC vaccine composition to the patient. Treating a patient with a cancer that
CC expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
CC comprises administering to the patient a vector encoding single chain
CC monoclonal antibody that comprises the variable domains of the heavy and
CC light chains of the monoclonal antibody that specifically binds to STEAP,
CC such that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02781) in figure 9A-9C
CC of the specification. However the present sequence lacks several amino
CC acids at its N-terminal end and has additional amino acids at its C-
CC terminal end when compared with the sequence shown in figure 9A-9C
XX
SQ Sequence 450 AA;

Query Match 73.9%; Score 1738; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No.1.1e-178;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 NMRINQYPESNAEYLASLFPDLSLYKGFNVSAWALQIGPKDASRQVYICSNNIQARQV 178
Db 1 NMRINQYPESNAEYLASLFPDLSLYKGFNVSAWALQIGPKDASRQVYICSNNIQARQV 60
QY 179 IELARQINFIPIIDIGSLSSAREIENLPLRLFTLWRGPPVVAISLAEFFLYSFVRDVIHP 238
Db 61 IELARQINFIPIIDIGSLSSAREIENLPLRLFTLWRGPPVVAISLAEFFLYSFVRDVIHP 120
QY 239 YARNQSDFYKIPIEIVNKTLPVATLTLVYLAGLAAAYQLYGTYRRFPPLLETW 298
Db 121 YARNQSDFYKIPIEIVNKTLPVATLTLVYLAGLAAAYQLYGTYRRFPPLLETW 180
QY 299 LQCRKQLGLISFFPAMVHVAYSLCPMRRSERYLFLNMAVQVYHANIENSMNEEVRRIE 358
Db 181 LQCRKQLGLISFFPAMVHVAYSLCPMRRSERYLFLNMAVQVYHANIENSMNEEVRRIE 240
QY 359 MYISFGIMSGLLSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFVLIYWKRA 418
Db 241 MYISFGIMSGLLSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFVLIYWKRA 300
QY 419 FEEYRYRFTPPNFVLALVLPISVILDLQLCRYPD 454
Db 301 FEEYRYRFTPPNFVLALVLPISVILDLQLCRYPD 336
RESULT 15
ABP62883
ID ABP62883 standard; protein; 1082 AA.
XX
AC ABP62883;
XX
DT 14-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 320.
XX
KW Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytosolic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
OS Homo sapiens.
XX
PN WO200218424-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027093.
XX
PR 01-SEP-2000; 2000US-00654935.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX
DR WPI; 2002-583321/62.
DR N-PSDB; ABQ93362.
XX
PT New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies.
XX
PS Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising one of
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the protein
CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (II) to (II).
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC (I) is useful for gene therapy of diseases and (II) can be used for
CC therapeutic treatment. Diseases that may be treated include wound healing
CC and tissue repair, burns, central nervous system disorders (e.g.
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC sclerosis, diabetes and allergies. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SO Sequence 1082 AA;

Query Match 58.0%; Score 1364.5; DB 5; Length 1082;
Best Local Similarity 83.2%; Pred. No. 1.1e-137;
Matches 272; Conservative 0; Mismatches 0; Indels 55; Gaps 1;

QY	120	MRINQYPSNAEYLA	SLFPD	SLIVKGFNV	SAMALQ	LGPKDASROY	YICSNNIQARQ	QVI	179
DB	1	MRINQYPSNAEYLA	SLFPD	SLIVKGFNV	SAMALQ	LGPKDASROY	YICSNNIQARQ	QVI	60
QY	180	ELARQLNFIPID	GLSSAREI	ENLPLRLFTL	MRGPNVAISL	ATFFFLYSFVR	DVTHPY	239	
DB	61	ELARQLNFIPID	GLSSAREI	ENLPLRLFTL	MRGPNVAISL	ATFFFLYSFVR	DVTHPY	120	
QY	240	ARNQOSDFYKIP	IEIVNKTLP	IVAITLSLVYLA	GLLAAYQLYG	TKYRRFP	PWLETWL	299	
DB	121	ARNQOSDFYKIP	IEIVNKTLP	IVAITLSLVYLA	GLLAAYQLYG	TKYRRFP	PWLETWL	180	
QY	300	QCRKQLGLSFF	PAMVHVA	YSLCLPMR	RSERYLFLN	MAYQQVHAN	IENSWNEE	EWRIEM	359
DB	181	QCRKQLGLSFF	PAMVHVA	YSLCLPMR	RSERYLFLN	MAYQQVHAN	IENSWNEE	EWRIEM	220
QY	360	YISFGIMSLGL	SLAVTSIP	SVSNALNWR	EFSTIO	STLGYVALL	ISTFVLI	YWKRAF	419
DB	221	-----	-----	-----	-----	QSTLGYVALL	ISTFVLI	YWKRAF	245
QY	420	EEEYRFFYTP	PNFVLA	LVLP	SI	VLDL	446		
DB	246	EEEYRFFYTP	PNFVLA	LVLP	SI	VLDL	272		

Search completed: March 1, 2004, 23:33:05
Job time : 102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 23:31:21 ; Search time 45 Seconds
(without alignments)
970.466 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGMGSPKLSLSTCLPN.....ALVLPISIVLIDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	6.8	198	2	D95285 conserved hypothet
2	159.5	6.8	211	2	AC2560 hypothetical prote
3	156.5	6.7	239	2	T50571 probable oxidoredu
4	147	6.3	213	2	H69400 conserved hypothet
5	142	6.0	212	2	D69361 conserved hypothet
6	142	6.0	224	2	T10120 F420-dependent MAD
7	133	5.7	223	2	D64487 hypothetical prote
8	121	5.1	191	2	F86826 hypothetical prote
9	120.5	5.1	232	2	A69131 conserved hypothet
10	119.5	5.1	216	2	T00121 hypothetical prote
11	116.5	5.0	242	2	G82642 conserved hypothet
12	116	4.9	198	2	AB3182 conserved hypothet
13	111.5	4.7	695	1	JN0898 folitropin recept
14	111.5	4.7	1228	2	S59681 probable membrane
15	110.5	4.7	222	2	B84410 hypothetical prote
16	110.5	4.7	694	2	JC4301 folitropin recept
17	109	4.6	320	2	T28379 ORF MSV218 hypothe
18	109	4.6	712	2	S50969 probable membrane
19	109	4.6	1184	2	H71436 hypothetical prote
20	109	4.6	1301	2	D85188 disease resistance
21	107.5	4.6	629	2	S60385 probable membrane
22	107.5	4.6	695	1	QRHUF folitropin recept
23	105	4.5	442	2	B64582 sodium- and chlori
24	105	4.5	604	2	TJ1042 hypothetical prote
25	104	4.4	396	1	C64907 chloramphenicol re
26	104	4.4	396	2	G90895 hypothetical prote
27	104	4.4	396	2	H85721 probable resistanc
28	103	4.4	465	2	S69915 sodium-phosphate t
29	102.5	4.4	314	2	AB2972 hypothetical prote

30	102.5	4.4	314	2	H98310 probable peptide A
31	102.5	4.4	1242	2	T39453 probable mra stab
32	102	4.3	220	2	AG3547 bicyclomycin resis
33	102	4.3	574	2	T41068 hypothetical prote
34	101	4.3	320	2	E71139 hypothetical prote
35	101	4.3	348	2	T12284 NADH2 dehydrogenas
36	101	4.3	420	2	F69144 O-antigen transpor
37	101	4.3	501	2	T02134 hypothetical prote
38	101	4.3	503	2	C86250 hypothetical prote
39	101	4.3	735	2	A83006 hypothetical prote
40	100.5	4.3	346	2	T11181 NADH2 dehydrogenas
41	100.5	4.3	1780	2	A85045 probable glucan sy
42	100	4.3	395	2	C71219 hypothetical prote
43	100	4.3	476	2	A28439 endonuclease Scea
44	100	4.3	714	2	AF2479 ABC transporter At
45	99.5	4.2	452	2	C71391 NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

conserved hypothetical protein Sma0349 [imported] - Sinorhizobium meliloti (strain 1021
D95285
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C/Accession: D95285
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Aboia, A.P.; Barloy-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A/Reference number: A95262; MUID:21396509; PMID:11481432
A/Accession: D95285
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-198 <KUR>
A/Cross-references: GB:AE006469; PIDN:AAK64846.1; PID:914523260; GSPDB:GNC00165
A/Experimental source: strain 1021, megaplasmid pSymA
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: Sma0349
A/Genome: plasmid

Query Match 6.8%; Score 160; DB 2; Length 198;
Best Local Similarity 26.7%; Pred. No. 2.1e-05;
Matches 55; Conservative 40; Mismatches 83; Indels 28; Gaps 8;

Qy	32	TVGVIGSGDPAKSLTIRLCGYHVI-GSRNPKFAS---EFFPHVDVTHEDALTKTN	87
Db	3	TYAIGAGAGSALAERFTAAQIPAIANSRGPASLSVTDRFGASVKAVELKDAL-QAD	61
Qy	88	IIFVAIHREHYTSLMDLRHLIV---GKILIDVSNMNRINQYP-----ESNAEYLASLFP	138
Db	62	VVILAV---PYDSIADIVTQSDWGQIVVDASNAIDFPAPKPRDLGRLSTEIVSELVP	118
Qy	139	DSLIVKGFNVSAWALQLGPK--DASRQVYICSNNIQARQVIELARQLNFIPIIDGSL	196
Db	119	GAKVVAFNTLPAAVLAADPKTGTSRVFLSGNHSDPANRQVAELISLGFAPVDGLTA	178
Qy	197	SAREIENLPLRLFTLMRGPVVAISL	222
Db	179	ASGPICQF-----GRPLVALNL	195

RESULT 2

AC2560
hypothetical protein alr8074 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga
C:/Species: Nostoc sp. PCC 7120
A:/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:/Accession: AC2560
R:/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:/Reference number: AB1807; MUID:21595285; PMID:11759840
A:/Accession: AC2560
A:/Status: Preliminary
A:/Molecule type: DNA
A:/Residues: 1-211 <KUR>
A:/Cross-references: GB:AP003603; PIDN:BAB77404.1; PID:g17134847; GSPDB:GN00182
A:/Experimental source: strain PCC 7120
C:/Genetics:
A:/Gene: alr8074
A:/Genome: plasmid

[illegible]

```

RESULT 3
T50571
probable oxidoreductase [imported] - Streptomyces coelicolor
C:/Species: Streptomyces coelicolor
C:/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
C:/Accession: T50571
R:/Redenbach, M.; Kieser, H.M.; Denapaité, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:/Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:/Reference number: Z20556; MUID:97000351; PMID:8843436
A:/Accession: T50571
A:/Status: preliminary; translated from GB/EMBL/DBJ
A:/Molecule type: DNA
A:/Residues: 1-239 <RED>
A:/Cross-references: EMBL:AL133220; PIDN:CAB61708.1
A:/Experimental source: strain A3(2)
C:/Genetics:
A:/Note: SCC75A.08c
C:/Superfamily: conserved hypothetical protein MJ1501

```

	Query Match	6.7%;	Score 156.5;	DB 2;	Length 239;
	Best Local Similarity	27.9%;	Pred. No. 5e-05;		
	Matches	61;	Conservative 41;	Mismatches 72;	Indels 45; Gaps 10;
QY	25 IKDARKVTGVVI-GSGDFAKSLTIRLRCGYHVVIGSR--NPKFASEFPFHVVDVTHHE	80			
Db	23 LPDVSGLVGVLGGTGPOEGCLAYRLAKAGQKVIVGSRAAERAAAAAEIGHGEVGADNA	82			
QY	81 DALTKNTIIFVAIHRE-HYTSIMDLRHLVGGKILIDVSNNMRINQ-----YPE--SNAE	131			
Db	83 ETARSDVVIIVAAVPWDGHGHTLESIRAEISGLLVDCVNPDIQFDKKGAYALKPEEGSAE	142			

```
QY      132 YLASLPFDSLIVKGFNVVSANALQ-----LGPKDASRQVYTICSNNIQARQCV   178
           ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      143 QAALLPDSRVAAAFHHLSAVLQDPEIDETDTVMVLGEERADAVEI-----VQA----   192
                                           ||| :||| :||| :||| :||| :|||
QY      179 IELARQLNFIP---IDLGLSLSAREIENLPLRLFTLMR    213
           ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      193 --LAGR---IPGMRGVFAGRLLRNAHQVESLVANLISVNR    226
```

RESULT 4
H69400
conserved hypothetical protein AF1209 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: H69400
R./Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A./Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A./Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A./Reference number: A69250; MUID:98049343; PMID:9389475
A./Accession: H69400
A./Status: preliminary; nucleic acid sequence not shown; translation not shown
A./Molecule type: DNA
A./Residues: 1-213 <KLE>
A./Cross-references: GB:AE001021; GB:AE000782; NID:g2689344; PIDN:AAB90038.1; PID:g26493
C./Superfamily: conserved hypothetical protein MJ1501

[illegible]

RESULT 5
D69361
conserved hypothetical protein AF0892 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: D69361
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
. ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: D69361
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-212 <KLE>
A/Cross-references: GB:AE001042; GB:AE000782; NID:g2689365; PIDN:AAE90348.1; PID:g264970
C/Superfamily: conserved hypothetical protein MJ1501


```

Query Match      6.0%; Score 142; DB 2; Length 212;
Best Local Similarity 25.4%; Pred. No. 0.0006;
Matches 49; Conservative 44; Mismatches 84; Indels 16; Gaps 6;

QY      37 GSGDFAKSLTIRLRICGYHVIVIGSRNPK---FASFFPHVDVT---HMEDALTKTNI 88
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      8 GTGNLKGGLRLPLATLGHEIIVGSSRREEKAKEAKAEYRIAGDASITGMKNEDAAEACDI 67

QY      89 IFVAIHREH-VTSIMDLRHLLVGKI---LIDVSNMNRINQY--PESNAEYLASLFPSDL 141
       :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      68 AVLTIPWEHAIDTARDLKNILREKIVSPLYEVSRGAKEFTYSSERSAELIAVEYLESEK 127

QY      142 IVKGFNVSAWALQLGPKDASROYVICSNNIQAROOVIELARQLNFI-PIDIGSLSSARE 200
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      128 VVSALHTTTPARFRANLDEKFDWDVPVCDDDESKKVMSLSIEIDGLRPDLADGPLNSRSL 187

QY      201 IENLPRLFTLMR 213
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      188 VESLTPLINIMR 200

```

RESULT 6
T10120
F420-dependent NADP reductase (EC 1.6.8.-) - Methanobacterium thermoautotrophicum (strain
C/Species: Methanobacterium thermoautotrophicum
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C/Accession: T10120
R/Berk, H.; Thauer, R.K.
FEBS Lett. 438, 124-126, 1998
A/Title: F420H2:NADP oxidoreductase from Methanobacterium thermoautotrophicum: Identific
A/Reference number: Z16959; MUID:99037734; PMID:9821972
A/Accession: T10120
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-224 <BER>
A/Cross-references: EMBL:Y17210
A/Experimental source: strain Marburg
C/Superfamily: conserved hypothetical protein MJ1501
;Keywords: oxidoreductase

	Query Match	6.0%;	Score 142;	DB 2;	length 224;	
	Best Local Similarity	29.2%;	Pred. No. 0.00064;			
	Matches	59;	Conservative	29;	Mismatches	84;
					Indels	30;
					Gaps	8;
QY	37 GSGDFAKSLTIRLRCGYHVIVIGSRNPKFASFEFFPHVVDVTHEDATK--TN-----					87
Db	8 GTGDQGLALRLALAGEEVIIGSRDAEKAVSAAQKVLEIAERDDLKVGKATNAEAEEA					67
QY	88 ---IIFVAIHREHYTSLWDLRHLVLGKILIDV-----SNMRINQYPESNAEYLA					134
Db	68 EVALITVPLQAQMAT-LGSVKKAIKGVLLIDATVPIDSCUGSAVRYYIDLWDGSAAEERA					126
QY	135 SLFPD--SLIVKGFNVVSAMALQ--LGPKDASROYVICSNNIQARQOVIELARQLNFT-P					189
Db	127 RFLEDQGSTRVAAAAFNNISASALLDITGPVDC--DCLIASDHRDALDLASELAEEKIDGVR					184
QY	190 IDLGSLSSAREIENLPRLFTL					211
Db	185 IDCGLLENARVIEKITPPLLINTL					206

RESULT 7
 D64487
 hypothetical protein MJ1501 - *Methanococcus jannaschii*
 C/Species: *Methanococcus jannaschii*
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C/Accession: D64487
 R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: D64487
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <BUL>
A;Cross-references: GB:L067591; GB:L77117; NID:g2826422; PIDN:AAB99514.1; PID:g1500389;
C;Genetics:
A;Map position: REV1473617-1472946
C;Superfamily: conserved hypothetical protein MJ1501

```
Query Match      5.7%; Score 133; DB 2; Length 223;  
Best Local Similarity   21.9%; Pred. No. 0.0033;  
Matches    46; Conservative     52; Mismatches   78; Indels   34; Gaps    7;
```

QY		37	GSGDFAKSTITRLIRCGYHVIGSRNPKFASFFPHVVDV-----THHEDALT	84
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db		8	GTGDQGFGIALRLAK-NNKIIISRKKKEAKEPAKKAKEILKGIEADITGLENDKDAK	66
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY		85	KTNIIIFVAIHREH-YTSLMDLRHLIVGKILIDV-----SNNMRINQYPE-SNAEYL	133
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db		67	EADVVLISLPYEYTLSTIKQLKEELKGKIIVSIGVPPLATAGDKPTRLLFPDDGSVAEMV	126
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY		134	ASLFPDSLIVKGFENVVSAMALQLGPKDASROYVICSNNIQARQVIETLARQLNFI-PIDL	192
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db		127	QNVLKESKVSAFQNVCHAVLELDNPVDCDILLVCGNDEEAKKVVIDLANQIIDGVRAIDC	186
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY		193	GSLSSAREIENLPRLFTLMRGPVVAISL	222
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db		187	GNLEKSRIIEAIT-----PLLGLINI	207

RESULT 8
F86826
Hypothetical protein ygfE [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C/Species: *Lactococcus lactis* subsp. *lactis*
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: F86826
R/Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: F86826
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-191 <STO>
A/Cross-references: GB:AE005176; PID:g12724621; PIDN:AAK05712.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: ygfE

```

Query Match          5.1%; Score 121; DB 2; Length 191;
Best Local Similarity 22.2%; Pred. No. 0.024;
Matches 47; Conservative 33; Mismatches 74; Indels 58; Gaps 7;

QY      32 TWVGIGSGDPAKSLTIRLIRCGYHAVVIGSRNPKFASEFFPHVDVTHEDALTKTN--- 87
        |||::|||::|
Db       3 TISIFGKGKMGKAI-----GDNFSSSVKNKYIILSNSKTELGEI 42

QY      88 -----IIFVAIHREHYTSLWDLRHLVGLIDVSNNMRINQYPE-----SNAEYL 133
        :||::||::|
Db       43 VWLAVPYAVALGIIEYSTD-----LQGKIIDIITNPVDFTFPSLLVPSDTSAALI 95

QY      134 ASLFPSDLLVKGFNVVSAMALQLGPDKASRQ----VYICSNNIQAQQVIETLARQLNFIP 189
        |:||:||||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       96 AKOLPNSMIVAKFN--TTFSDDLTAIKKVANEHQTVTLASDSQEAKETIIKALENSGLST 153

QY      190 IDLGSLSSAREIENTLPRLFTL-----WRG 214
        ::|||::|::|::|::|::|::|
Db       154 LDAGSLKRARELEAIGFLQITTLASEKISWDG 185

```

RESULT 9
A69131

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Delta H)
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: A69131
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Tiwari, N. K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional annotation
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: A69131
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-232 <MTH>
A/Cross-references: GB:AE000811; GB:AE000666; NID:G2621287; PIDN:AAB84754.1; PID:G2621289
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH248
A/Start codon: TTG
C/Superfamily: conserved hypothetical protein MJ1501

	Query Match	5.1%;	Score 120.5;	DB 2;	Length 232;
	Best Local Similarity	23.8%;	Pred. No. 0.034;		
	Matches	49;	Conservative 42;	Mismatches 90;	Indels 25; Gaps 6;
QY	30 KVTWGVIGSGDFAKSLTIRLCRGYHVIGSRNPKFASFEPPHVDVTHHE-----	80			
Db	8 KIAV-IGTGDDQLGLALNFVAAGEEVIIGSRDAEKASKASKYLEIAGRDDISVEGATN	66			
QY	81 -DALTKNIIFVAIH-REHYTSIMDLRHLVKGILIDVS-----NNMRINQYPESN	129			
Db	67 PDAASADVVLTVPDAQMTLASIRDQRDKVLIDAYVIDSCIGSAVRYYIDLMEGS	126			
QY	130 AEYLASF--PDSLIVKGFNVVSAMALOGLPKDASRQYITCSNNIQARQQVIELARQLN	186			
Db	127 AAEARAARFLREQGTREVAAFNNNISASALLEVSEPVDCCLVASDHRDALEVAAPLAEKID	186			
QY	187 FI-PIDLGSLSSAREIEMLPLRLFTL	211			
Db	187 GVRAIECGGLENAARIETKITPLLINT	212			

RESULT 10
T00121
hypothetical protein 8 - Leptospira interrogans

C|Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C|Accession: T00121
R|Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998
A|Title: Physical and genetic maps of the *Leptospira interrogans* serovar icterohaemorrhagiae
A|Reference number: Z14115; MUID:98332717; PMID:9666070
A|Accession: T00121
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-216 <TAK>
A|Cross-references: EMBL:AB010203; NID:g2780763; PIDN:BAA24368.1; PID:g2780765
A|Experimental source: strain Ictero No.1; substrain icterohaemorrhagiae

Query Match	5.1%;	Score 119.5;	DB 2;	Length 216;
Best Local Similarity	23.7%;	Pred. No. 0.037;		
Matches 47;	Conservative 37;	Mismatches 91;	Indels 23;	Gaps 6;

```
QY      33 VGVIGSGDFAKSLTIRLRGCGYHVIVIGSRNPKFASEFFPHV---VDVTHEDALTKTNI 89
      :::|||      :::|||      :::|||      :::|||
Db      6 IGILGSGIVGQTIANGLFELRYGAELVKTGRDFGKLKDWLAKAGASIGSFSEANFGEII 65

QY      90 FVAIHREHYTSLMDLRHL--LVGKILIDVSNMR-----INQYPESNAEYLA 136
      :::|||      :::|||      :::|||      :::|||
Db      66 VLCSKGSVASEVLTLSGIDSLNGKTTIIDTNPISSEIPONGVINFFTSYNESLMEKLOKQ 125

QY      137 FPDSLIVKGFNVVSAMALQIGP--KDSARQVYICSNNICARQOVIELARQNLNFIPIDIGS 194
```

Db 126 APKANFVKCFSSVGS-GLMVPQLKGEKPSMFCIGNDDSSKKQIKELIDTFCWDTEDMGK 184

QY 195 LSSAREIENLPRLFTLW 212

Db 185 VEAPRAIEPLCI---LW 198

RESULT 11
G82642

conserved hypothetical protein XF1737 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: G82642
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: G82642

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-242 <SIM>
A;Cross-references: GB:AE003997; GB:AE003849; NID:g9106805; PIDN:AAF84546.1; GSPDB:GN00
A;Experimental source: strain 9a5c
R;Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasae
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1737

Query Match	5.0%;	Score 116.5;	DB 2;	Length 242;
Best Local Similarity	24.0%;	Pred. No. 0.075;		
Matches 46;	Conservative 40;	Mismatches 85;	Indels 21;	Gaps 7;

[illegible]

RESULT 12
AB3182
conserved hypothetical protein Atu5183 [imported] - *Agrobacterium tumefaciens* (strain C
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AB3182
R;Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; Mccl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB3182
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45872.1; PID:gl7743615; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atus183
A:Genome: plasmid

Query Match 4.9%; Score 116; DB 2; Length 198;
Best Local Similarity 22.7%; Pred. No. 0.063;
Matches 50; Conservative 41; Mismatches 81; Indels 48; Gaps 10;

QY 31 VTVGVIGSGDFPAKSLTIRLRIGYHVI-GSRNPKFASFFPHV--VDVTHHEDALTXT 86
Db 1 MTVGIIIGAGNIGGAFATAGKAGIEAVIANSRGPESLTALVSKLGSTIRAGVPEAAQA 60
QY 87 NIFVAIHREHYTSLWD-----LRHLVG-KILIDVSNMNRINOY-----PESNAEYLA 134
Db 61 -IVLVAVP-----WSKIPGALAGLNFGRVIDANNSIEAPLYRPADLGRTSTDIPT 112
QY 135 SLFPDSLIVKGFNVSAWALQLGPKDAS-----RQVYICSNNIQARQVIELARQL 185
Db 113 ALVPGARVVKAFN-----HLTPKQLSGDPHSEGGRRVLFYSGDDMRKAKEVGALIDRI 165
QY 186 NFPIPIDLSLSSAREIENLPRLFTLMRGPVVAISLATF 225
Db 166 GFPGIDLGGLPVGSQLQFP-----GGPLPALNLVKF 197

RESULT 13
JN0898

foliitropin receptor precursor - crab-eating macaque
N:Alternate names: follicle-stimulating hormone receptor (FSHR)
C/Species: Macaca fascicularis (crab-eating macaque)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: JN0898; S36452
R/Gromoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.
Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993
A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of th
A:Reference number: JN0898; MUID:94071854; PMID:7504463
A:Accession: JN0898
A:Molecule type: mRNA
A:Residues: 1-695 <GRO>
A:Cross-references: EMBL:X74454; NID:g396801; PIDN:CAAS2463.1; PID:g396802
A>Note: the authors translated the codon AGT for residue 488 as Arg
C/Function:
A:Description: receptor that mediates the biochemical effects of foliitropin
C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein;
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-695/Product: foliitropin receptor #status predicted <PRH>
F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:367-387/Domain: transmembrane #status predicted <TM1>
F:399-421/Domain: transmembrane #status predicted <TM2>
F:444-465/Domain: transmembrane #status predicted <TM3>
F:486-508/Domain: transmembrane #status predicted <TM4>
F:529-550/Domain: transmembrane #status predicted <TM5>
F:574-597/Domain: transmembrane #status predicted <TM6>
F:609-630/Domain: transmembrane #status predicted <TM7>
F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.71;
Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 HEDALTKNTIIFVAIHREHYTSLMDLRHLVG-----KILIDVSNMNR 121
Db 98 HEIRIERKANNL-LYINPEAFQNPMLRYLLISNTGIKLPDVHKIHSFQKVLDDIGNIN 156
QY 122 INQYPESNAEYLASLFPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIEL 181
Db 157 IH-----TIERNSPVGLSPESVILWL-----NKGCIQEIHNCA----- 189
QY 182 ARQLNFIPIIDLGLSSAREIENLPRLFTLMRGPVVAISLATFFLYSFVRDVIHPYAR 241
Db 190 ---FNGQLDELNTLSDNNNLEELPNDVFHGASGPVILDISRTIRHSLPSYGLLENLKURA 246
QY 242 NQSDPFIKPIPIEIVNKTLPYVAITLLSLV----- 271
Db 247 RSTYNLKLIP-----SLKLVALLMEASLTYPSSHCCAPANWRROISELHPICNKSILROEV 301
QY 272 -----LAGLLAAVQLYYGTYRRF-----PMLLETWLQCKROLG 306
Db 302 DYMOTIRGORSSLAEDNESSYRGFPDMTYAEFDYDLCEVVDVTCSPKPDAFNPCEDIIG 361
QY 307 -----LISFFAMVHA-----YSLCLPMRBSERYFLNMAVQOVHANIENS 348
Db 362 YNILRVLIWFISILAITGNIIVLVTLTTSQKLTVP-----RLMCNLAFAD----- 408
QY 349 WNEEYVRIEMYISFGIMSLGLLSLAVTSIPSVSNALNRE-----FSFIQSTLG 399
Db 409 -----LCIGIYLLIASVDIHTKSQYHNAVLDWQTGAGCDAAAGFTVPASBELS 456
QY 400 ---YVALLISTFH-----VLIYGMKRAFEERYRFTYPPNFV-LAL 436
Db 457 VYTLTAITLERWHTITHAMQLDCKVHRHAASVWVGWTFAPFAALFPIFGISSYMKYSI 516
QY 437 VLPSIVILLDLQL 449
Db 517 CLPMDIDSPLSQL 529

RESULT 14
S59681

probable membrane protein YPL012w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YPL5w; hypothetical protein YP8132.01
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C/Accession: S59681; S52519
R/Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; W
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A:Reference number: S59677
A:Accession: S59681
A:Molecule type: DNA
A:Residues: 1-1228 <HAL>
A:Cross-references: EMBL:U33335; NID:g965076; PID:g965081; MIPS:YPL012w
R/Baddock, K.; Churcher, C.
submitted to the EMBL Data Library, February 1995
A:Reference number: S52519
A:Accession: S52519
A:Molecule type: DNA
A:Residues: 220-1228 <BAD>
A:Cross-references: EMBL:Z48483; NID:g683777; PID:g683778; MIPS:YPL012w
A:Experimental source: strain AB972
C/Genetics:
A:Gene: SGD:RRP12
A:Cross-references: SGD:S0005933
A:Map position: 16L
C/Keywords: transmembrane protein
F:125-141/Domain: transmembrane #status predicted <TM1>
F:478-494/Domain: transmembrane #status predicted <TM2>
F:720-736/Domain: transmembrane #status predicted <TM3>

Query Match	4.7%;	Score 111.5;	DB 2;	Length 1228;
Best Local Similarity	19.9%;	Pred. No. 1.5;		
Matches 111;	Conservative 71;	Mismatches 152;	Indels 225;	Gaps 28;

```

QY 11 KSLSETCLPENGINGIKDARK-----VTGVIGSGDEAKSLTIRLI 50
Db 292 KTMABTTISSGLAENKXLRVLDTIFALKPSNVDLLTKSWIAVAVIKMSTVATHQPLKAL 351
QY 51 R-----CGYHVVIGSRNPKF--ASEFFPHVVDVTHEDAL-----83
Db 352 RKIPGVFHIMCTY---LASETPEVYQAASQCLISLSESVKDDLLLYTPSVDEKVEKXND 408
QY 84 -----TKTNIIFFAIREHYTSLMDLRHLVGLKILIDVSNMRINQYPESNAEYLAJLF 137
Db 409 EIIISOAKTFIDFLSIRYSHCS-----REIL--KILVAAFNKFRY---RSNPHFLKSL- 456
QY 138 PDSLIVKGFNVVSAMALQLPKASRQVYICSNNTQARQVTEIARQNLFI-----188
Db 457 -----KIVDTWRYN-----EEQFMDLRNEIELVYGASISAMG 488
QY 189 -----PIDLGSLSAREIENLPLRLFTLWRGPVY---VAISLATFFFLYSFVRDVI 236
Db 489 PEMILAEAPLNLDPSSER-----PGR--ANLLPLIRDYTKNANLATF-----QNEL 533
QY 237 HPYARNQOSDFYKPIELEIVN-----KTLB-----IVAITLJSL 269
Db 534 APYIKSFQSKFDKYPEESIQLRVQOTIVDQIWSITLPRCELPMDLRESFTDEFASELSSL 593
QY 270 VYL-----AGLLAAAYQLYYGK-----YRFP-----PWEITWLQCRK 303
Db 594 LYSEVELRTTICHALKVLAEBSNVSYAESSSHVYLLQRFPISEAQKNIERYIST----K 648
QY 304 QLGLLSFFAMTVHAYSLCLPMRRS-----ERYL-----FLMAYQOVHANIENSWE 351
Db 649 STNLLAVLFNV---YTQTTPNARSYILETTIDQYLKITSKEDLEKTFNNVCGILKNMSNE 704
QY 352 EEWRIEMYSISGINSGLJLSL--AVTISIPSVSNALNWRREFSFIQSTGLYVALLISTFH 409
Db 705 ESSGNVNKEKKKPOLATLDDLITICMITYLPVSSYSALFSMFS-----LTVNSAD 754
QY 410 VLIYGWKRAFEERYRYFT 428
Db 755 ALIQ-KRA-----YRIIT 766

```

RESULT 15
B84410
hypothetical protein Vng2607c [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C/Accession: B84410
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84410
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <STO>
A;Cross-references: GB:AE004437; NID:g10581987; PIDN:AAG20646.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG2607C
C;Superfamily: conserved hypothetical protein MJ1501

Query Match	4.7%;	Score 110.5;	DB 2;	Length 222;
Best Local Similarity	24.3%;	Pred. No. 0.2;		
Matches 49;	Conservative 34;	Mismatches 94;	Indels 25;	Gaps 6;

37 GSGDEAKSLTIRLCGYH-VVIGSRNPKFASEFFPHVDVT-----HEDAL 83

Db 8 GTGDDIGAGLALRWATDSDHDIVIGSRDPEKARETAAYEDTLADQGVDRKLTGFANEMAA 67

QY 84 TKTNIIFVAIHREHYTSLW-----DLRHLLVGKILIDVSNMNRINQYPSNAE---YL 133

Db 68 DRADVVAAPAYHTVDWGAVALRDLADTLVISPAVGIASGEHGLHYNPPSAGSYTALV 127

QY 134 ASLFPDSL-IVKGFNVSAWALQLGPKDASRQVYICSNNIQARQOVIELARQLNFI-PID 191

Db 128 ADAAPDGVVDVGAFFHNLAAADRLLADLTLELDADTLVGVNDEGARTVAELADDTIGLRALD 187

QY 192 LGSLSARBIENLPRLFTLWR 213

Db 188 AGPEVENAAEVESLTPLLINLAR 209

Search completed: March 2, 2004, 06:12:46
Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 20:23:21 ; Search time 45 Seconds
(without alignments)
520.849 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESTSMWGPSKLSLSETCLPN.....ALVPSIVLIDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	38.3	173	4	US-09-323-873A-8 Sequence 8, Appli
2	736	31.3	141	3	US-09-083-521-1 Sequence 1, Appli
3	717	30.5	339	4	US-09-323-873A-2 Sequence 2, Appli
4	717	30.5	339	4	US-09-685-166A-879 Sequence 879, App
5	144.5	6.1	227	4	US-09-655-270A-15 Sequence 15, Appli
6	144.5	6.1	227	4	US-09-651-941-17 Sequence 17, Appli
7	144.5	6.1	227	4	US-09-955-597-17 Sequence 17, Appli
8	107.5	4.6	695	1	US-08-487-886-2 Sequence 2, Appli
9	107.5	4.6	695	3	US-08-482-855-2 Sequence 2, Appli
10	107.5	4.6	695	4	US-08-474-986-2 GENERAL INFORMA
11	107	4.6	34	4	US-09-323-873A-20 Sequence 20, Appli
12	102.5	4.4	940	4	US-09-328-352-8165 Sequence 8165, Ap
13	101	4.3	476	3	US-09-316-083-3 Sequence 3, Appli
14	101	4.3	476	4	US-09-933-700-3 Sequence 3, Appli
15	101	4.3	724	4	US-09-252-991A-21494 Sequence 21494, A
16	99	4.2	365	4	US-09-170-496D-118 Sequence 118, App
17	99	4.2	365	4	US-09-170-496D-226 Sequence 226, App
18	99	4.2	365	4	US-09-364-425B-27 Sequence 27, Appli
19	99	4.2	531	2	US-08-724-974A-2 Sequence 2, Appli
20	97.5	4.1	390	3	US-08-460-576-2 Sequence 2, Appli
21	97	4.1	692	3	US-07-757-342D-6 Sequence 6, Appli
22	97	4.1	692	4	US-09-461-657B-6 Sequence 6, Appli
23	96	4.1	299	4	US-09-903-456-61 Sequence 61, Appli
24	96	4.1	1309	4	US-09-975-413A-10 Sequence 10, Appli
25	94.5	4.0	288	4	US-09-107-532A-6009 Sequence 6009, Ap
26	94.5	4.0	296	4	US-09-540-236-2922 Sequence 2922, Ap
27	94.5	4.0	345	4	US-09-489-039A-10740 Sequence 10740, A

28	94	4.0	284	4	US-09-903-456-83 Sequence 83, Appli
29	94	4.0	292	4	US-09-903-456-35 Sequence 35, Appli
30	94	4.0	293	4	US-09-903-456-45 Sequence 45, Appli
31	93	4.0	365	2	US-08-724-974A-3 Sequence 3, Appli
32	93	4.0	498	4	US-09-107-532A-7077 Sequence 7077, Ap
33	92.5	3.9	405	4	US-09-489-039A-9411 Sequence 9411, Ap
34	91.5	3.9	407	4	US-09-328-352-5605 Sequence 5605, Ap
35	91.5	3.9	495	4	US-08-985-343-2 Sequence 2, Appli
36	91	3.9	211	4	US-09-252-991A-23822 Sequence 23822, A
37	91	3.9	440	4	US-09-634-238-281 Sequence 281, App
38	91	3.9	550	4	US-09-543-681A-6026 Sequence 6026, Ap
39	91	3.9	808	4	US-09-134-001C-3105 Sequence 3105, Ap
40	90.5	3.8	340	4	US-09-543-681A-7503 Sequence 7503, Ap
41	90.5	3.8	388	4	US-09-107-532A-6381 Sequence 6381, Ap
42	90	3.8	500	4	US-09-178-093B-26 Sequence 26, Appli
43	90	3.8	970	4	US-09-795-927-7 Sequence 7, Appli
44	89.5	3.8	495	1	US-07-841-997A-2 Sequence 2, Appli
45	89.5	3.8	495	1	US-08-290-301-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-323-873A-8
Sequence 8, Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Leong
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129.16USU2
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 173
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 DFYKPIPIVNTLPIVAITLISLVYLAGLAAAYQLYGKRRFPFWLWTWLCRKQL 305
DB 1 DFYKPIPIVNTLPIVAITLISLVYLAGLAAAYQLYGKRRFPFWLWTWLCRKQL 60
QY 306 GLISFFPAMVHVAISLCLPMRSEERYLFLNMAQQVHANINENSWNEBEVRIEMYSFGI 365
DB 61 GLISFFPAMVHVAISLCLPMRSEERYLFLNMAQQVHANINENSWNEBEVRIEMYSFGI 120
QY 366 MSGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRA 418
DB 121 MSGLISLLAVTSIPSVSNALNWFESFIQSTLGYVALLISTFHVLIYWKRA 173

RESULT 2
US-09-083-521-1
Sequence 1, Application US/09083521
Patent No. 6048970
GENERAL INFORMATION:

```

; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT10
; CLONE: 1691243
;
US-09-083-521-1

Query Match          31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 MHHVAYSLCLPMRSERYLFLNMAYQOVHANIENSWNEEVRRIEMYISFGIMSLGLSL 373
DB 1 MHHVAYSLCLPMRSERYLFLNMAYQOVHANIENSWNEEVRRIEMYISFGIMSLGLSL 60

QY 374 LAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGWKRAFEERYRFTPNFV 433
DB 61 LAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGWKRAFEERYRFTPNFV 120

QY 434 LALVLPSTIVLDLQLCRYPD 454
DB 121 LALVLPSTIVLDLQLCRYPD 141

RESULT 3
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
```

```

; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
;
US-09-323-873A-2

Query Match          30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVRDVIHPYARNOQSDFFYKPIPIEIVNKTLPVATITL 267
DB 67 LFPQWHPPIKIAIIASLTFPLYTLREVIVHPLATSHQOYFYKPIPIVINKVLPVMSITLL 126

QY 268 SLVYLAGLAAAYOLYGTKYRRFPPEWLETWLQCRKQLGLSFFPAMVHAYSLCLPMR 327
DB 127 ALVYLPGVIAIVQLHNGTKYKFPWIDKMLTRKQFGLSFFPVLHAIYSLSPMR 186

QY 328 SERYLFLNMAYQOVHANIENSWNEEVRRIEMYISFGIMSLGLSLAVTSIPSVSNALN 387
DB 187 SYRKLLNMAYQOVQONKEDAWIEHDVMEIYVSLGIVGLAIALAVTSIPSVSDSLT 246

QY 388 WREFSFIQSTLGYVALLISTFHVLIYGWKRAFEERYRFTPNFVLAIVLSIVIL 444
DB 247 WREFHYIQSKLGIIVSLGLTIALIFANNKWDIKQFVWYTPPTFMIAVFLPIVILI 303

RESULT 4
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-685-166A-879

Query Match          30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVRDVIHPYARNOQSDFFYKPIPIEIVNKTLPVATITL 267
```


[illegible]

RESULT 5

```

US-09-655-270A-15
; Sequence 15, Application US/09655270A
; Patent No. 6329151
;
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic MR
; FILE REFERENCE: BC1011 US NA
; CURRENT APPLICATION NUMBER: US/09/655,270A
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; PRIOR APPLICATION NUMBER: 60/152,542
; PRIOR FILING DATE: 1999-September-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
;
; LENGTH: 227
;
; TYPE: PRF
;
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

```

Query Match	6.1%;	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches	60;	Conservative	33;	Mismatches 88;
			Indels	33;
			Gaps	9;

```

QY      25 IKDKARKYIVGIVIGSGDFAKSLTIRLCGYHVVIGSRNPKFASEFFPH-----VVD 75
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      1 MKSSKIAYVG--GTGPQGGKGLAYRFAAGMPVIVIGSRSAEAEAEALLEVRRRAGDGAVVS 58
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

QY      76 VTHHEDALTKENIIFVAI----HREHYTSLMDLRHLVLVKILLI-----DVSNNMRIN 123
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      59 AADNNSAADCPILLLVFPYDGHRELVS---ELAPIFAGLTVSCVNPLGFDKSGAYGLD 115
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

QY      124 QYPESNAEYIASLFPDSLIVKGFNVSA---WALQLGPKDASRQVYIGSNNIQARQVTE 180
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      116 VEEGSAEQRLDLVPGATVVAAFHHLSAVNLMEME-GP--LPEDVLVCGDDRSADQEVAR 172
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

QY      181 LARQLNFIIP-IDGLSSAREIENIPLRLFTLWR 213
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      173 LAVAITGRPGIDGALRVARQLLEPLTAVLINVR 206
      :  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

```

RESULT 6

```

US-09-651-941-17
; Sequence 17, Application US/09651941
; Patent No. 6355470
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/651,941
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-651-941-17

```

Query Match	6.1%;	Score 144.5;	DB 4;	Length 227;
-------------	-------	--------------	-------	-------------

Best Local Similarity 28.0%; Pred. No. 6.4e-07;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;

```
QY      25 IKDARKVTGVIGSGDEFAKSLTIRLCRGHVIVIGSRNPKFASEFFPH-----VVD 75
        :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 MKSSKIAVG--GTCPQGKLAYRFAAGMPEVIGSRSARAEAALEVRRRAGDGAVVS 58
```

QY 76 VTHEDALTKNIIFFVAI---HREHYTSLMDLRHLVGKILI-----DVSNNMRIN 123
 : | :
Db 59 AADNASAAADCPITILLVPPYDGRELVS---ELAPIFAGKLVSQVNPLGFDKSGAYGLD 115

```

QY      124 QYPSNAEYLASLFPDLSLYKGNVSA---WALQLGPKDASRQVYICSNIIQARQVIE 180
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      116 VEEGSAEQLRDLPFGATVVAAFHHLISAVNLMHEH-GR--LPEDVLYCGDPRSADKDEVAR 172

```

```
QY      181 LARQLNFIPI-DIGSLSSAREIENLPURLFTLMR 213
        ||| : ||| : ||| : | : |
DB      173 LAVAITGRPGIDGALLRVARQLEPPLTAVLINVR 206
```

RESULT 7

```

US-09-955-597-17
; Sequence 17, Application US/09955597
; Patent No. 6461856
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BCI022 US NA
; CURRENT APPLICATION NUMBER: US/09/955,597
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PR1
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-955-597-17

```

Query Match	6.1%;	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches	60;	Conservative	33;	Mismatches 88;
			Indels	33;
			Gaps	9;

Oy 25 IKDARKVTGVIIGSGDFAKSLTIRLCGHVIVIGSRNPKPASEFFPH-----VVD 75
 ::| | :| | :| | | | :| |
Db 1 MKSSKIAVG--GTGPQKGGLAYRFAAAGENVVIGSRAERAEEALEVRRLGDGAVVS 58

```
OY      76 VTHHEDALTKINII FVAI --- HREHYTSLMDLRHLVGKILI ----- DVSNNMRIN 123
           : | : : : ||| : : : | : :
Db     59 AADNAAADCPILLV PYDGHRELVS -- ELAPIFAGKLTVSCVNPPLGPDKS GAYGLD 115
```

```

OY      124 QYPESNAEYTLASTLPDLSLVKGFNVSA---WALQLGPKDASRQVYTCISNNIQARQVIE 180
      ||| ||| : : : ||| : ||| : : : |||
Db      116 VEEGSAEQLRDLVPGATVYAAFHLLSAVNLMEHE-GP--LPEDVLVCGDDRSKAKDEVAR 172

```

```
QY      181 LARQLNFIIP-IDIGSLSSARELENIPLRLFTLMR   213
        || : ||| |::| : |
Db      173 LAVAITGRPGIDGALLRVARQLEPLTAVLINVR    206
```

RESULT 8

US-08-487-886-2

Sequence 2, Application US/08487886
Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massil via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448

IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2
Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;
QY 79 HEDALTKENIIFVAIHREHYTSLMDRLHLVG-----KILIDVSNMR 121
DB 98 HEIRIEKANML-LYINPEAFQNLPLQYLLISNTGKILPDVHKIHSLOKVLDIQDNIN 156
QY 122 INQPSNAEYLAFLPDSLIVKGFNVSAWALQLPKDASROYVICSNNIQARQVTEL 181
DB 157 IH-----TIERNSFVGLSFESVILWL---NKNGIQEIHNCA----- 189
QY 182 ARQLNFIPIDLGSLSSAREIENLPLRFTLMRGPVVAISLATFFLYSFVRDVIHPYAR 241
DB 190 ---FNGTQDELNLSDNNLLEELPNDVFHGASGPVILDISRTIRIHSLSYGLNKKLRA 246
QY 242 NQSDFYKPIEIVNKTLPVATLLSLV----- 271
DB 247 RSTYNLKKLP-----TLEKLVALMEASLTYPHSCAFANWRQISELHPICNKSILRQE 300
QY 272 -----LAGLLAAAYQLYGTKYRF-----PPMLBTWLQCRKQL 305
DB 301 VDYMTQTRGSSSLAEDNESSYSGFDMTYTEFDYDLCEVVDYTCSPKPAFNPCEIDIM 360
QY 306 G-----LSFFFAVHVA-----YSLCLPMRGRERYLFLMAYQQVHANIN 347
DB 361 GYNILRVLIWFISILAITGNITIVVILTSQYKLTVP-----RFLMCNLAIFAD----- 408
QY 348 SWNEEVRRIEMTISFGIMSLGLSLAVTSIPSVSNALNRE-----FSFIQSTL 398
DB 409 -----LCIGIYLLIASVDIHTKSQYHNYAIDMQTGAGCDAAGFTVFASGL 455
QY 399 G---YVALLISTFH-----VLITGWKRAFEERYRFPYTPPNFV-LA 435
DB 456 SVYTLTATITERWHTITHAMQLDCYQLRHAASVVMGMWIFAFALPPIFGISSYMKVS 515
QY 436 LVLPSVILDLQL 449
DB 516 ICLPMDIDSPISQL 529


```

RESULT 9
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massil via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

```

[illegible]

Db 516 ICLPMDIDSPISQL 529

RESULT 10

US-08-474-986-2

GENERAL INFORMATION:

APPLICANT: Kelton, Christie Ann

Schweickhardt, Rene Lynn

Cheng, Shirley Vui Yen

Mugent, No. 637271leen Patrice

TITLE OF INVENTION: Human Follicle Stimulating

Hormone Receptor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Stephan P. Williams,

Ares-Serono, Inc.

STREET: Exchange Place, 37th floor

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density

COMPUTER: IBM PS/2, model 55 SX

OPERATING SYSTEM: MS-DOS version 4.0

SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474, 986

FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/670, 085

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Stephan P.

REGISTRATION NUMBER: 28546

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 723-1300

TELEFAX: (617) 723-8923

LOCATION: 614 to 678

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-474-986-2

Query Match 4.6%; Score 107.5; DB 4; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

79 HEDALTNTNIIFAIHEHYTSLMDLRHLVG-----KILIDVSNMNR 121
98 HEIRIEKANML-LYINPEAFQNLPLQYLLISNTGIKHLPDVHKIHSLOKVLDDIODNIN 156
122 INQYBESNAEYLASLPDLSLVKGFNVSAWALQLPKDSARQVYICSNNIQARQVIEL 181
157 IH-----TIERNSEVGLSFESVILWL-----NKGIGQEIHNCA----- 189

182 ARQNFIPIDGLSSAREIENLPLRLFTLWGPVVAISLATFFLYSFVRDVITHPYAR 241
190 ---FNGTQDLDELNLSDNNNLELLPNDVFHGAAGPYIIDISRTRIHSLPSYGLNLKGLRA 246
242 NQOSDFYKIPRIEIVNKL-PIVAITLISLV----- 271
247 RSTYNLKKLP-----TLEKLVALMEASLTYPHCCAFANWRROISELHPICNKSILRQE 300

272 -----LAGLIAAAYQLYGTKYRRF-----PPWLETWLQCRKQL 305
301 VDYMTQTRGQRSSLAEDNESSYSRGFDMTYTEFDYDLCEVVDVTCSPREDAFNPCEDIM 360
306 G-----LISFFAMVHA-----YSUCLPMRSERYLFLNAYQOVHANIEN 347
361 GYNILRVLIWFISITAITGNIIVLITTSQYKLTVP-----RFLMCNLAFAD----- 408

QY 348 SWNEEWEVRIENYISFGIMSLGLSLAATSIPSVSNALNWR-----FSFIQSTL 398
Db 409 -----LCIGIYLLIASVDIHTKSQYHNYAIDWQTAGCDAGFTVFASEL 455

QY 399 G---YVALLISTFH-----VLIGWKRAFEERYRFTYPPNFV-LA 435

Db 456 SVYTLTATLERWHTITHAMQLDCKVQLHHAASVWVMGMIFAFPAALFPIFGISSYMKVS 515

QY 436 LVLPSIVILDLQL 449

Db 516 ICLPMDIDSPISQL 529

RESULT 11

US-09-323-873A-20

Sequence 20, Application US/09323873A

Patent No. 6329503

GENERAL INFORMATION:

APPLICANT: Daniel E. Afar

APPLICANT: Rene S. Hubert

APPLICANT: Kahan Leong

APPLICANT: Arthur B. Raitano

APPLICANT: Douglas C. Saffran

APPLICANT: Steve Chappell Mitchell

TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS

TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF

FILE REFERENCE: 129.16USU2

CURRENT APPLICATION NUMBER: US/09/323, 873A

CURRENT FILING DATE: 1999-06-01

PRIOR APPLICATION NUMBER: 60/087, 520

PRIOR FILING DATE: 1998-06-01

PRIOR APPLICATION NUMBER: 60/091, 183

PRIOR FILING DATE: 1998-06-30

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 20

LENGTH: 34

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: STEAP-1 PEPTIDE

US-09-323-873A-20

Query Match 4.6%; Score 107; DB 4; Length 34;
Best Local Similarity 58.8%; Pred. No. 0.0003;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 326 RRSERYLFLNWAYQOVHANIENSWNEEWEVRIEM 359
Db 1 RRSRYKLLNWAYQOVQNKEDAWIEHDVWRMEI 34

RESULT 12

US-09-328-352-8165

Sequence 8165, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328, 352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 8165

LENGTH: 940

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-8165

Query Match 4.4%; Score 102.5; DB 4; Length 940;
Best Local Similarity 19.5%; Pred. No. 0.15;
Matches 71; Conservative 74; Mismatches 146; Indels 73; Gaps 17;


```
QY      88 IIFVAIHREHYTSLMDLRHLLVGKILIDVSNMRINQYPESNAEYLASL-----FPDSL I 142
      262 VVFLTLVESHHGCV--ITSLTTASILFTILLAITVFLSLKQQAIIYLAIALGMAVAAPIV 319
QY      143 VKGF--NVYSAMALQLGPKDASRQVYICSNNIQARQOVIELAROLN---FIPIDLSLS 196
      320 IPQYRPDVVFLFS-----YVLVINLAVAAVNIQPMKILINQIAFFATMFIGSA 368
QY      197 SAREIENLPRLFTL-WRGPVVAISLATEFFELYSFVRDVIHPYARNQOSDFYKIPLEIV 255
      369 IAFYAE--PAKFDLDM---ILWLHIALFIWLSVRYSQNTSRVSEHEKQEGIRLP---- 418
QY      256 NKTLPVATLLSLVYLAGLAAAQLYYGTK-----YRRFPWLJETWLQCRK 303
      419 ----FLDVGILFNVPVLGFTLHAYLVHESQTALTI GAAVLAGTYAVLTFWIK---KTHP 471
QY      304 QLGLSFFPAMVHVA-YSLCLPMRSEERYLLFLNMAQQOVHANINENSWNEEVRIMETIS 362
      472 QLSVLAKSFFILLAVAFALLIFLAKGHWTAIGVAAQGTALIV--WGVTERTYRLSRYI- 527
QY      363 FGIMSLGLSLAVTSIPSVSNALNW--REFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
      528 -GVILVLSLALFYQV-----WANEEFPTLSTSIYAIQFISAFYLLQYNSK---E 575
QY      421 EEEY 424
      576 QRYF 579
```

RESULT 13
US-09-316-083-3
Sequence 3, Application US/09316083A
Patent No. 6280942

GENERAL INFORMATION:
APPLICANT: The Institute of Physical and Chemical Research
TITLE OF INVENTION: Endonuclease
FILE REFERENCE: PH-651
CURRENT APPLICATION NUMBER: US/09/316,083A
CURRENT FILING DATE: 1999-05-20
EARLIER APPLICATION NUMBER: JP98/141861
EARLIER FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 476
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-316-083-3

Query Match 4.3%; Score 101; DB 3; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

```
QY      62 NPKFASEFFPHVVDVTHHEDALTKNI---IFVAIHREHYT-----SLMD 103
      254 NPYFVNAFSINI-----KTNLAKEKIFTNINYKLYSDYKINQINNHIPIYNYLK 302
QY      104 LRHLVKGKILIDVSNMRINQYPESNAEYLASL--PDSLIVKGFNVVSAMALQLGPKDA 161
      303 INNKLPIKNIMDIKNMYWLAGFYADGSFLSSMYPKDTLLEKNM----- 347
QY      162 SRQVYICSNNIQARQOVIELARQINFIPIDLSLSAREIENLPRLFTLWRGPVVAIS 221
      348 -RPSYVIS-QVETRKEIYLIG3---SFDL-SISNVKKGVRKLDKDFKLFRTTDELNR 400
QY      222 LATFFELYSFVRDVIHPYARNQOSDFYKIP---IEIVNKTLPVAVITLLSLVYLAGLLA 277
      401 -----FIYYF--DKFLPLHDNKQENYIKFRFNTFIKSYNNNRVFGVLSE--YINNIKI 451
QY      278 AAYQLYYGTKY 288
      452 DNYDYYYYNKY 462
```

RESULT 14
US-09-933-700-3
Sequence 3, Application US/09933700
Patent No. 6528296
GENERAL INFORMATION:
APPLICANT: The Institute of Physical and Chemical Research
TITLE OF INVENTION: Endonuclease
FILE REFERENCE: PH-651
CURRENT APPLICATION NUMBER: US/09/933,700
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 09/316,083
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 476
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-933-700-3

Query Match 4.3%; Score 101; DB 4; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

```
QY      62 NPKFASEFFPHVVDVTHHEDALTKNI---IFVAIHREHYT-----SLMD 103
      254 NPYFVNAFSINI-----KTNLAKEKIFTNINYKLYSDYKINQINNHIPIYNYLK 302
QY      104 LRHLVKGKILIDVSNMRINQYPESNAEYLASL--PDSLIVKGFNVVSAMALQLGPKDA 161
      303 INNKLPIKNIMDIKNMYWLAGFYADGSFLSSMYPKDTLLEKNM----- 347
QY      162 SRQVYICSNNIQARQOVIELARQINFIPIDLSLSAREIENLPRLFTLWRGPVVAIS 221
      348 -RPSYVIS-QVETRKEIYLIG3---SFDL-SISNVKKGVRKLDKDFKLFRTTDELNR 400
QY      222 LATFFELYSFVRDVIHPYARNQOSDFYKIP---IEIVNKTLPVAVITLLSLVYLAGLLA 277
      401 -----FIYYF--DKFLPLHDNKQENYIKFRFNTFIKSYNNNRVFGVLSE--YINNIKI 451
QY      278 AAYQLYYGTKY 288
      452 DNYDYYYYNKY 462
```

RESULT 15
US-09-252-991A-21494
Sequence 21494, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21494
LENGTH: 724
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21494

Query Match 4.3%; Score 101; DB 4; Length 724;
Best Local Similarity 20.5%; Pred. No. 0.14;
Matches 72; Conservative 57; Mismatches 105; Indels 118; Gaps 16;

```
QY      132 YLASLFPDSLIVKGFNVVSAMALQLGPKDASRQVYICSNNTQA-----RQOVIELARQIN 186
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      298 YLSVVLPPDSL-SKTLAMWMAAYVLVCGITLFSA---LCVTISLSLGGPHRORALDILRRQA 352
QY      187 FIPIDL-GSLSSAREIENLPRLFTLRGPVVVAISLA-----TFPFLYSFVRDYIHP 238
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      353 FRPLWLIGSLALGEVAHDPRLIAGLGHTSICLSTLANASAALEFALFVMFRFRPIAH- 411
QY      239 YARNQQ-----SDPYKIPI-----EIVNK 257
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      412 LIRNQLERLRKRSLHDLVOLVGSLMEVEVLVLVGISLFATFEVSAGDSSSALRRALVCA 471
QY      258 TLPIVAITLLSLVYLAGLIAAAYQLYYGTKYRFRPFWLETWLQCRKQLGLTSFFAWHEV 317
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      472 VLAVAMTVIGLI-----RRRSRYGAGPRRSAPYIEQ-----LOSFGYTLLHI 515
QY      318 AYSLCLPMRSEERYLFLNMAYQQVHANIENTSWNEEBEWRIEM-YISFGIM----- 366
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      516 FVVVL-----FFIEVALRWGMSELIR-YAEGBEGEQISMKVVSFGTTLVLVAMLIIWL 564
QY      367 -----SLGL-----LSLLAVTSPSVSNALNWREBSFIQTULGYVALLIS 406
        ||::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      565 TDTAIQHSLGLGGKSREPNTRALTMPLIRNLV-----PATIAVIALLIVA 608
```

```
Search completed: March 1, 2004, 23:40:58
Job time : 48 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 19:51:19 ; Search time 40 Seconds
(without alignments)
590.996 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMWGPSKLSLSTCLPN.....ALVLPISIVLLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	30.5	339	1 STEA_HUMAN	Q9uhes homo sapien
2	133	5.7	223	1 FARE_METUA	O58896 methanococc
3	120.5	5.1	232	1 FARE_METTH	O26350 methanobact
4	111.5	4.7	695	1 FSHR_MACFA	P32212 macaca fasc
5	109.5	4.7	206	1 Y538_PASMU	Q9cn97 pasteurella
6	109	4.6	693	1 FSHR_CHICK	P79763 gallus gall
7	109	4.6	712	1 FRE6_YEAST	Q12473 saccharomyc
8	107.5	4.6	629	1 FRE7_YEAST	Q12333 saccharomyc
9	107.5	4.6	695	1 FSHR_HUMAN	P23945 homo sapien
10	104.5	4.4	695	1 FSHR_PIG	P49059 sus scrofa
11	104	4.4	396	1 SOTB_ECO57	P58529 escherichia
12	104	4.4	396	1 SOTB_ECOLI	P31122 escherichia
13	103	4.4	465	1 NPT1_MOUSE	Q61983 mus musculu
14	102.5	4.4	391	1 NUCC_NEPOI	Q9tkv6 nephroselmi
15	102	4.3	220	1 Y304_BRUME	Q8ydt3 brucella me
16	102	4.3	220	1 Y301_BRUST	Q8fvs9 brucella su
17	100	4.3	476	1 RF3_SACBA	P05512 saccharomyc
18	99.5	4.2	452	1 NU4M_BRALA	O79421 branchiosto
19	99	4.2	365	1 SPRI_HUMAN	Q15743 homo sapien
20	98	4.2	3411	1 POLG_YEFV1	P03314 y genome po
21	98	4.2	3411	1 POLG_YEFV2	P19901 y genome po
22	97.5	4.1	268	1 YD49_AQUAE	O67364 aquifex aeo
23	97.5	4.1	695	1 FSHR_BOVIN	P35376 bos taurus
24	97	4.1	218	1 YG46_XANAC	Q8p1y9 xanthomonas
25	97	4.1	692	1 FSHR_RAT	P20395 rattus norv
26	96.5	4.1	686	1 FRE1_YEAST	P32791 saccharomyc
27	96	4.1	261	1 PHSC_ECOLI	P77409 escherichia
28	95	4.0	202	1 YAJ1_PSEAE	Q91v45 pseudomonas
29	95	4.0	393	1 NUCC_ARATH	P56753 arabidopsis
30	95	4.0	393	1 NUCC_SPIOL	Q9m315 spinacia ol
31	95	4.0	692	1 FSHR_MOUSE	P35378 mus musculu
32	94.5	4.0	492	1 SECY_CYAPA	P25014 cyanophora
33	94.5	4.0	695	1 FSHR_SHEEP	P35379 ovis aries

34	93.5	4.0	365	1 SPRI_MOUSE	Q8bfq3 mus musculu
35	93.5	4.0	452	1 NU4M_BRALA	O47423 branchiosto
36	93.5	4.0	499	1 GSHR_PLAF7	O15770 plasmodium
37	93.5	4.0	530	1 AAAL_MOUSE	Q9jmh8 mus musculu
38	93.5	4.0	1233	1 VL1_REOVD	P15024 reovirus (t
39	93	4.0	388	1 HMC3_DESVH	P33390 desulfovibr
40	93	4.0	459	1 NU4M_BOVIN	P03910 bos taurus
41	93	4.0	521	1 YT25_CAEEL	Q10934 caenorhabdi
42	92.5	3.9	490	1 GTRI_CHICK	P46896 gallus gall
43	92.5	3.9	500	1 NU4C_OENHO	P58419 oenothera h
44	92.5	3.9	883	1 YHL6_YEAST	P38781 saccharomyc
45	92.5	3.9	1769	1 YUK9_YEAST	P42945 saccharomyc

ALIGNMENTS

RESULT 1	ID	STEA_HUMAN	STANDARD;	PRT;	339 AA.
AC	Q9UHE8; Q95034;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Six transmembrane epithelial antigen of prostate.				
GN	STEAP OR STEAP1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=20056277; PubMed=10586738;				
RX	Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,				
RA	Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,				
RA	Jakobovits A., Saffran D.C., Afar D.E.H.;				
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in				
RT	human prostate tumors.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).				
RN	[2]	SEQUENCE FROM N.A.			
RP	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;				
RA	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
RL	[3]	SEQUENCE FROM N.A.			
RP	TISSUE=Skin;				
RC	MEDLINE=22388257; PubMed=12477932;				
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.;				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahay J., Hellton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalhus D.E.,				
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- TISSUE SPECIFICITY: Highly expressed in prostate tumors.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF186249; AAF17479.1; -.
DR EMBL; AC005053; AAC79150.1; ALT_INT.
DR EMBL; AC004969; AAD15620.2; -.
DR EMBL; BC011802; AAH1802.1; -.
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Antigen.
FT TRANSMEM 71
FT TRANSMEM 91
FT TRANSMEM 119
FT TRANSMEM 164
FT TRANSMEM 218
FT TRANSMEM 258
FT TRANSMEM 291
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 1; Length 339;
Best Local Similarity 54.9%; Pred. No. 6e-48;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVAISLATPFFLYSTVDYIHPARNQSDFYKIPFIEVNTLPVATILL 267
DB 67 LFPQWHLPIKTAIAIASLTFLYTLREVIHPLATSHQYFYKIPILVINKVLPWVSTILL 126
QY 268 SLVYLAGLLAAAYQLYGTQRRPPEMETWLQCRKQLGLSPFPAMVHVAYSTLEMER 327
DB 127 ALVYLPGVIAIVQLHNGTKYKPPHMLDKMLTRKQGLSPFAVLHAIYSTLEMER 186
QY 328 SERYLFLNMAVQOVHANIEVSWNEEVRLEMYISFGIMSLGLSLAVTSIPVSALN 387
DB 187 SYRYKLINMAVQOVQNKEDAWIEHDVMEIYVSLGIVGLAIIALLAVTSIPVSALN 246
QY 388 WREFSFIQSTLGYVALLISTFHVLIYWKRAFEERYRFPENFVALVPSIVL 444
DB 247 WREFHYIQSKLGIIVSLIGTIIHALIFANWKWIDIKQFVWYTPPTFMIAVFLPIVLI 303

RESULT 2

F4RE_METJA STANDARD; PRT; 223 AA.
ID F4RE_METJA
AC 058896;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN M01501.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the

CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -!- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U67591; AAB99514.1; -.
DR PIR; D64487; D64487.
DR TIGR; M01501; -.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 223 AA; 24068 MW; 2370BBD6F5BBD3D9 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;
Best Local Similarity 21.9%; Pred. No. 0.003;
Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDPAKSLTIRLRGYNHVGSRNPKPASEFFPHVVDV-----THHEALT 84
DB 8 GTGDQGFGLALRLAK-NNKIIIGSRKKEAEAKAKEILKRGIEADIIGLENKDAK 66
QY 85 KTNIIFAIHEH-YTSLMDLRHLVGLIDV-----SNMNRINQYPE-SNAEYL 133
DB 67 EGDVILSLPYEYTLSTIKQLKEELKGIKIVSIGVPLATAIGDXPTRLFPDGSVAEMV 126
QY 134 ASLPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIELARQLNFI-PIDL 192
DB 127 QNVLKESKVSFAQNVCHAVLEDDNPDVDILVCGNDEBAKVYIDLANQIDGVRAIDC 186
QY 193 GSLSSAREIENPLRLFTLWRGPVVAISL 222
DB 187 GNLEKSRITFAIT-----PLIGLNI 207

RESULT 3

F4RE_METTH STANDARD; PRT; 232 AA.
ID F4RE_METTH
AC 026350;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN MTH248.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).


```
CC -!- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -!- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000811; AAB84754.1; -.
DR PIR; A69131; A69131.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
KM Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 232 AA; 24539 MW; A1CE60ABC8474296 CRC64;

Query Match
Best Local Similarity 23.8%; Score 120.5; DB 1; Length 232;
Matches 49; Conservative 42; Mismatches 90; Indels 25; Gaps 6;

QY 30 KATVGVIGSGDFAKSLTIRLCRGYHVIGSRNPKASEFPFHVVDVTHHE----- 80
DQ 8 KIAV-IGGTGDOGLALRFVAGEEVTIGSRDAEKASKASKVLEIAGRDISVEGATN 66
QY 81 -DALTKNTIIFVAIH-REHYTSLMDLRHLVGLKILVDS-----NNNRINQYPSN 129
DQ 67 PDAAASADVVLTVPLQAQWVTLASIRDQVRDKVLIDATVPIDSCIGSAVYIDLWEGS 126
QY 130 AEYLASLF---PDSLIVKGFNVVSAWALQLGPKASRQVYICSNNIQARQVIELARQLN 186
DQ 127 AAEBAARFLREQGTRVAAAFNNISASALLEVSEPVDCDLVASDHRDALEVAELAEKID 186
QY 187 FI-PIDGLSLGSAREIENLPLRLFTL 211
DQ 187 GVRAIECGLENNARIIEKITPLLNL 212

RESULT 4
FSHR_MACFA STANDARD; PRT; 695 AA.
ID FSHR_MACFA
AC P32212;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Folicle stimulating hormone receptor precursor (FSH-R) (Folllitropin
DE receptor).
GN FSHR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OC NCBI_TaxID=9541;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94071854; Pubmed=7504463;
RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
RT "Molecular cloning of the testicular follicle stimulating hormone
RT receptor of the non human primate Macaca fascicularis and
RT identification of multiple transcripts in the testis.";
RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
CC -!- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylylate cyclase.
```

```
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74454; CA52463.1; -.
DR PIR; JN0898; JN0898.
DR HGSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCRHDOPSN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT TRANSMEM 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CRC64;

Query Match
Best Local Similarity 4.7%; Score 111.5; DB 1; Length 695;
Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 HEDALTKNTIIFVAIHREHYTSLMDLRHLV-----KILDVSNMR 121
DQ 98 HEIRIEKANNL-LYNPEAFQNLPLRYLLISNTGIKLDPVHKHSFGKVLIDIQDNIN 156
QY 122 INQYPSNAEYLASLFPDSLIVKGFNVVSAWALQLGPKASRQVYICSNNIQARQVIEL 181
DQ 157 IH-----TIENSFVGLSFESVILML-----NKGIOEIHNCA----- 189
QY 182 ARQLNFIPIIDGLSLGSAREIENLPLRLFTLRGQVVAVVAISLATFFLYSPVVDVHPYAR 241
DQ 190 ---FNGTQDELNLSDNNNLELPNDVFHGASGPVILDISRTIRHSLPSGLENLKGLRA 246
```

```

QY 242 NQSDPYKPIPIVNTLPPIVATLLSLVY----- 271
DB 247 RSTYNLKLTP-----SLKXVALMEASLTYPHSCAFANWRQISELHPICNKSILRQEV 301
QY 272 -----LAGLLAAAYQLYGTYYRRF-----PWLFTWLQCRKQLG 306
DB 302 DYMGTQRGQRSLSAEDNESSYSGPDMTYAEFDYDLQNEVDYTCSPKPDAFNPCEIDL 361
QY 307 -----ILSFFPAMVHA-----YSLCLPMRSEERYFLNMAQQVHANIEHS 348
DB 362 YNLRVLIWIFISILAITGNILVLTLSQYKLTVP-----RFLMCNLAFA----- 408
QY 349 WNEEYVWRIEMYISFGIMSLGLSLAVTSIPSVSNALNWE-----FSFIQSTLG 399
DB 409 -----LCIGIYLLILASVDIHTKSQYHNYAIDMGTGAGCDAAAGFTVFASELS 456
QY 400 ---YVALLISTFH-----VLIYGMKRAFEERYRFTPPNFV-LAL 436
DB 457 VYTLTAITLERMHTITHAMQLDCVHVHHAASVMVMGNIFAFAALPFIIGISSYMKVSI 516
QY 437 VLPSTVILDLQL 449
DB 517 CLPMDIDSPISQL 529

```

RESULT 5

```

Y538_PASMU
ID Y538_PASMU STANDARD; PRT; 206 AA.
AC Q9CN97;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li J.L., Paustian M.L., Whitlam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006089; AAK02622.1; -.
DR HAMAP; MF_01207; -.
DR InterPro; IPR007916; UPF0191.
DR Pfam; PF05252; UPF0191; 1.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 7 26 POTENTIAL.
FT TRANSMEM 41 63 POTENTIAL.
FT TRANSMEM 76 98 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT TRANSMEM 142 164 POTENTIAL.
FT TRANSMEM 169 191 POTENTIAL.
SQ SEQUENCE 206 AA; 23653 MW; 2F277729D7F11194 CRC64;

```

Query Match 4.7%; Score 109.5; DB 1; Length 206;
 Best Local Similarity 23.8%; Pred. No. 0.18;
 Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

```

QY 228 LYSFVADVTH-----PYA-----RNQSDPYKPIPIVNTLPPIVATLLSLVYLAGL 275
DB 1 MLSLFRITIIHVCCLGPAWMLAVLLSGDESQLGADPIKETQHFLGFSALTILLMFIIG- 59
QY 276 LAAAYQLYGTYYRRFPWLFTWLQCRKQLGLSFFPAMVHAYSCLPMRSEERYFLN 335
DB 60 -----KYFYLLK-----QPQLQV---LRRALGLMAWFYVLHV-YA-----YLALE 96
QY 336 MAYQVHANIEHSWNEEYVWRIEMYISFGIMSLGLSLAVTSIPSVSNALNWEFSFIQ 395
DB 97 LGYD-----FSLFVQELVNR--GYLLIGAIAFLITLMALSSWSYLK--LKMGMWFFYL 146
QY 396 STLGYVALLISTFHVLIYGMKRAFEERYRFTPPNFVIALVLPSTVILDLQLC 450
DB 147 HQLGTYALLIGAIIH--YVW-----SVKNVTFSSWL--YLILSIMILC 184

```

RESULT 6

```

FSHR_CHICK
ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor.";
RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgham J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT of cFSH-R messenger ribonucleic acid in the ovary.";
RL Biol. Reprod. 55:1055-1062(1996).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylylate cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D87871; BAAL3487.1; -.
DR EMBL; U51097; AAC60030.1; -.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.

```



```

DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCRHHODPSN.
DR SMART; SM00013; LRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 693
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT TRANSMEM 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 693
FT REPEAT 45 68
FT REPEAT 69 93
FT REPEAT 95 118
FT REPEAT 119 143
FT REPEAT 168 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 47 47
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 268 268
FT CONFLICT 4 4
FT CONFLICT 88 88
FT CONFLICT 140 140
FT CONFLICT 174 174
FT CONFLICT 191 191
FT CONFLICT 329 329
SQ SEQUENCE 693 AA; 78697 MW; 46F98699635A1BEC CRC64;

Query Match 4.6%; Score 109; DB 1; Length 693;
Best Local Similarity 17.2%; Pred. No. 0.87;
Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;

QY 18 LPNG-INGIKDARKVTWG-----VIGSGDFAKSLTIRLRGCGYHVVIGSRNPKFASEFFP 71
DB 61 IPKGAFGLHLEKIEISQNDALIEIGNVFSSL-----PKL----- 97

QY 72 HVDVTHEDALTKTNIFVAIHREHYTSLMDLRHLVG-----KILI 114
DB 98 -----HEIRIEKANKL-WKIDQDAFOHLPRLRYLLISNTGLSFLPVVHKVHSFOKVL 149

QY 115 DYSNNMRINQYPSNAEYLAFLPDSLIVKGFNVASAMALQLGPKDASROYICSNNTQA 174
DB 150 DVQDNHIRTIERNTFEGLS--ESVILR-----LNKNGIQE 184

QY 175 RQGVIELARQLNFIPIDLGSLSSAREIENPLRLFTLMRGVVAISLATFPFLYSFVRD 234
DB 185 IKD-----HAFNGTCLDELNLSNDYNLEKLPKVFQGAIGFVLDISRTISFLPSHGLE 239

QY 235 VTHPYARNQOSDFYKIPETIVNKTLPVATITLSLVYLAGL----- 275
DB 240 FIKKLIRARSTYKLCPL--DVNKRSLIEANFTYPSHCCAFTRKTKONTERTYPICSMSPA 297

QY 276 -----LAAAYQLYGYTKY--RRF-----PPWLETWLOCR 302
DB 298 KODLGEQTGKRKRHRSAEDYISHYGRFGEVNEFDYGLCNEVVDVFCSPKPDAPNCE 357

QY 303 KOLG-----LLSFFFAVHVA-----YSLCLPMRRSERYLELMAVYQOVHAN 344

```

```

DB 358 DIMGYNVLRLWIFINILAITGNTVLLILLISSQYKLTVP-----RFILMNLAFAD----- 408
QY 345 IENSWNEEVRRIEMYISFGIMSLGLLSLAVTSIPSVSNALNRE-----FSFIQ 395
DB 409 -----LCIGTYLFLIASVDIQTKSRYYNYALDMQTGACNAGFTTVA 452
QY 396 STLGVALLISTFH-----VLIYGNKRAF----- 419
DB 453 SELSVYTLVTITLERWHTTITAMQLNRKRLRHAVIIMVFGMFAFTVALLPIFGISSYM 512
QY 420 -----EEEYRFTPPNFYALVLPISVILLDLQLCRY 452
DB 513 KVSICLPMHIEPTFSQAYV--IFLLVNLVLAFTVICICITCY 553

RESULT 7
FREQ6 YEAST STANDARD; PRT; 712 AA.
ID FREQ6 YEAST
AC 012473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)
DE (Ferric-chelate reductase 6).
GN FREQ6 OR YIL051C OR I0593.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansgorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Eutian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moesti D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Pottellette D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels P., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Han J., Hobeisel J.D.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RT Nature 387:87-90(1997).
RL CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRE / CYB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z47973; CA88006.1; -.
DR EMBL; Z73156; CA97503.1; -.
DR PIR; S50969; S50969.
DR Germonline; 142046; -.
DR SGD; S0003974; FREQ6.
DR InterPro; IPR002916; Ferric_reduct.
DR Pfam; PF01794; Ferric_reduct; 1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW FAD; NAD; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17
FT CHAIN 18 712
FT NP_BIND 493 499 FAD (Potential).

```

```

FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 478 498 POTENTIAL.
FT TRANSMEM 550 570 POTENTIAL.
FT DOMAIN 553 556 POLY-LEU.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;

```

Query Match 4.6%; Score 109; DB 1; Length 712;

Best Local Similarity 18.8%; Pred. No. 0.9; Mismatches 119; Indels 142; Gaps 17;

Matches 72; Conservative 51; Mismatches 119; Indels 142; Gaps 17;

```

QY 88 IIFVAIHREHYTSL-----WDLRLHVLGKILIDVSNMNRINOYPESNAEYLASL 136
DB 183 IIAVEFHMSHYNGNRLAFASRFVNYIRGHVLPFTFLVD---KXANHFKEFLNVEVFTGL 238
QY 137 FPDSLIVKGFNVWSAWALQLGPKDASRQVYICSNNIQARQVIELARQLNFIPIDLGLS 196
DB 239 MPNSL-----EAW----- 246
QY 197 SAREIENLPLRLFTLMRGPVVAISLAFPFY-YSFVRDVIHPYARNQOSDFYKIPIEIV 255
DB 247 -----IIFGYTLANIFLSISYIIDPYNLIFNSHLSQFTRL--LA 284
QY 256 NKTLPVATLTLSTLYLAGLLAAAYQLYVGTKYRRFPPLMETWLQCRKQLGLLSFFFAMV 315
DB 285 DRS-GILAFTOFPLITFTARNSLFELTGVKENSF-----ISFHKVIGRIMVLANATI 336
QY 316 H-VAYSLCLPMRSERYLFLNMAVQOVHANINENSWNEEVRIMYISFGIMSLGLLSL 374
DB 337 HSLSYSL-----FAINNAFK-----ISNK-----QLYMKREGIASITVLCVL 373
QY 375 AVTSIPSVSNALNWRSEFISQTLGYVALLI---STFVLIY-GWKRAFE-----EE 422
DB 374 LVLSLGIYRK---RHYEFPLYTHIILALFFYCCMQHVKIFNGWKENTVYSLLIWGLEK 429
QY 423 YRFPYTPNFVALVLPISIVILL 446
DB 430 LFRIMN---ILOFRFPKATLILN 449

```

RESULT 8

PRE7_YEAST

ID PRE7_YEAST STANDARD; PRT; 629 AA.

AC Q12333;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Ferric reductase transmembrane component 7 (EC 1.16.1.7) (Ferric-

DE chelate reductase 7).

GN PRE7 OR YOL152W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / FY1679;

RX MEDLINE=96132030; PubMed=8553699;

RA Casamayor A., Aldea M., Casas C., Herrero E., Gamo F.J.,

RA Lafuente M.J., Gancedo C., Arino J.;

RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast

RL chromosome XV containing seven new open reading frames.";

RL Yeast 11:1281-1288(1995).

```

CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRE / CYBB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z48239; CAA88276.1; -.
CC EMBL; Z74894; CAA99174.1; -.
CC PIR; S60385; S60385.
CC GerMOnline; 143574; -.
CC SGD; S0005512; FRE7.
CC DR InterPro; IPR002916; Ferric_reduct.
CC DR Pfam; PF01794; Ferric_reduct_1.
CC KM Oxidoreductase; Electron transport; Transmembrane; Iron transport;
CC FAD; NAD; Glycoprotein; Multigene family.
CC NP BIND 369 375 FAD (POTENTIAL).
CC FT TRANSMEM 46 66 POTENTIAL.
CC FT TRANSMEM 108 128 POTENTIAL.
CC FT TRANSMEM 168 188 POTENTIAL.
CC FT TRANSMEM 195 215 POTENTIAL.
CC FT TRANSMEM 238 258 POTENTIAL.
CC FT TRANSMEM 266 286 POTENTIAL.
CC FT TRANSMEM 293 313 POTENTIAL.
CC FT TRANSMEM 422 442 POTENTIAL.
CC FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 629 AA; 71996 MW; 2384480B9289C16F CRC64;

```

Query Match 4.6%; Score 107.5; DB 1; Length 629;

Best Local Similarity 21.7%; Pred. No. 1; Mismatches 77; Indels 83; Gaps 17;

Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

```

QY 214 GPVVVAISLAFPFYSEVRDVIHPYARNQOSDFYKIPIEIVNKTLPVATLTLSTLY-L 272
DB 123 GTFVLVMAATLTYLLYCFVP--HPFYR-PCAGFGSPPLSV--RAGIMASLVPFVFSL 175
QY 273 AG-----LAAAYQLYVGTKYRRFPPLMETWLQCRKQLGLLSFFFAMVVAVSLCLPM 325
DB 176 SGKINIVGMLVGLSYE-----KINIVHGW-----ASILCLFFSWVHV-----IPF 215
QY 326 RRSERYLFLNMAVQOVHANINENSWNEEVRIMYISFGIMSLGLSLAVTSIP----- 380
DB 216 LROARH---EGYERH-----QRWKASDWMR-----SGVPILFLNLMWLSLPIARRH 262
QY 381 --SVSNALNWRSEFISQTLGYVALLISTFHY-----LIYWKRAFEHEEYR- 425
DB 263 FYEIFLQLHW-----LAVGFYISLF--YHYPELNSHMYLVATIVVW---FAQLFYRL 311
QY 426 ----FYTPNFVALVLPISIVILL 444
DB 312 AVKGYLRPGRSFMASTIANVSIV 334

```

RESULT 9

FSHR_HUMAN

ID FSHR_HUMAN STANDARD; PRT; 695 AA.

AC P23945; Q16225;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin

DE receptor).

GN FSHR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;
 [1]
 RA SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.
 RC TISSUE=Ovary;
 RX MEDLINE=91222171; PubMed=1709010;
 RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;
 RT "Cloning and sequencing of human FSH receptor cDNA.";
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93246012; PubMed=1301382;
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzeja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Ahara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermann T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Dankbar B., Gudermann T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RN [6]
 RP SEQUENCE OF 286-695 FROM N.A., AND VARIANT SER-680.
 RX MEDLINE=95000244; PubMed=7916967;
 RA Gromoll J., Ried T., Holtgreve-Grez H., Nieschlag E., Gudermann T.;
 RT "Localization of the human FSH receptor to chromosome 2p21 using a
 RT genomic probe comprising exon 10.";
 RL J. Mol. Endocrinol. 12:265-271(1994).
 RN [7]
 RP 3D-STRUCTURE MODELING OF 49-228.
 RX MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and the nature of hormone-receptor interactions.";
 RL Structure 3:1341-1353(1995).
 RN [8]
 RP VARIANTS ALA-307; ARG-524 AND SER-680.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [9]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 RN [10]
 RP VARIANTS ALA-307 AND SER-680.
 RX MEDLINE=22054685; PubMed=12059813;

RA Asatiani K., Gromoll J., Eckardstein S.V., Zitzmann M., Nieschlag E.,
 RA Simoni M.;
 RT "Distribution and function of FSH receptor genetic variants in normal
 RT men.";
 RL Andrologia 34:172-176(2002).
 RN [11]
 RP VARIANT OHSS ILE-449.
 RX MEDLINE=22812036; PubMed=12930927;
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,
 RA de Poncheville L., Chaplot S., Savagner F., Crove A., Mathieu E.,
 RA Lahlou N., Descamps P., Misrahi M.;
 RT "A chorionic gonadotropin-sensitive mutation in the
 RT follicle-stimulating hormone receptor as a cause of familial
 RT gestational spontaneous ovarian hyperstimulation syndrome.";
 RL New Engl. J. Med. 349:753-759(2003).
 RN [12]
 RP VARIANT OHSS ASN-567.
 RX MEDLINE=22812037; PubMed=12930928;
 RA Smits G., Olatunbosun O., Delbaere A., Pierson R., Vassart G.,
 RA Coetagsliola S.;
 RT "Ovarian hyperstimulation syndrome due to a mutation in the
 RT follicle-stimulating hormone receptor.";
 RL New Engl. J. Med. 349:760-766(2003).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of this receptor is mediated by G proteins which activate
 CC adenylylate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P23945-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P23945-2; Sequence=VSP_001953, VSP_001954;
 CC -1- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.
 CC -1- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation
 CC syndrome (OHSS) [MIM:608115]. OHSS is a disorder which occurs
 CC either spontaneously or most often as an iatrogenic complication
 CC of ovarian stimulation treatments for in vitro fertilization. The
 CC clinical manifestations vary from abdominal distention and
 CC discomfort to potentially life-threatening, massive ovarian
 CC enlargement and capillary leak with fluid sequestration.
 CC pathologic features of this syndrome include the presence of
 CC multiple serous and hemorrhagic follicular cysts lined by
 CC luteinized cells, a condition called hyperreactio luteinalis.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC FSH/LSH/TSR subfamily.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M65085; AAA52477.1; -;
 CC EMBL; S59900; AAB26480.1; -;
 CC EMBL; M95489; AAA52478.1; -;
 CC EMBL; X68044; CAA48179.1; -;
 CC EMBL; S73199; AAB32071.1; -;
 CC EMBL; S73526; AAB32225.1; -;
 CC PIR; I57661; QRTUFT.
 CC PDB; 1XUN; 15-MAY-97.
 CC Genew; HGNC:3969; FSHR.
 CC MIM; 136435; -;
 CC MIM; 608115; -;
 CC GO; GO:0016021; C:integral to membrane; TAS.
 CC GO; GO:0005886; F:follicle stimulating hormone receptor activity; TAS.
 CC GO; GO:0004963; F:follicle stimulating hormone receptor activity; TAS.
 CC GO; GO:0007292; P:female gamete generation; TAS.
 CC GO; GO:000585; P:female gonad development; TAS.
 CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
KW 3D-structure; Polymorphism; Disease mutation.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT TRANSMEM 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
FT VARSPLIC 224 285
/FTID=VSP_001953.

Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best local similarity 17.0%; Pred. No. 1.1; Mismatches 148; Indels 185; Gaps 17;
Matches 84; Conservative 77;

QY 79 HEDALTKTNIIFVAIHREHYTSLWDRLHLLVG-----KILIDVSNMNR 121
DB 98 HEIRIEKANNL-LYINPEAFQNLPLQYLLISNTGIKHLDPVHKIHSLOKVLDDQDNIN 156
QY 122 INQYPSNAEYLASLPDSLIVKGFNVSAWALQPKDASROYVICSNNIQARQVIEL 181
DB 157 IH-----TIERNSFVGLSFESVILWL---NKGIOEIHNC----- 189
QY 182 ARQLNFIPIIDGLSSAREIENLPLRLFTLWRGPVVVAISLATEFFLYSFVRDVIHPYAR 241
DB 190 ---FNGTQDELNLSDNNNLLELPNDYFHGASGPVILDISRTRHSLPSYLENLKRLA 246
QY 242 NQOSDFYKPIPIIVNKTLP-PIVATITLSLVY----- 271
DB 247 RSTYNLKKLP-----TLKXVALMEASLTYPHSCAFANWRQISELHPICNKSILRQE 300
QY 272 -----LAGLLAAAYQLYGTYKRRF-----PWLLETWLQCRKQL 305
DB 301 VDYMTQTRGQSSSLAEDNESSYRGFDMTYTEFDYDLQNEVVDVTCSPKPDAFNCEIDIM 360
QY 306 G-----LISFEFAMVHA-----YSLCLPMRSEERYLFLNMAYYQVHANIEN 347
DB 361 GYNILRLVLIWFISILAITGNITIVVILTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SWNEEEVVRLEMYISFGIMSLGLSLAVALTIPSVSNALNWR-----FSFTQSTL 398

DB 409 -----LCIGIYLLIASVDIHTKSYHNVAIDMQTAGCDAAGFTVFASEL 455
QY 399 G---YVALLISTFH-----VLITGWKRAFEERYRFTYTPNFV-LA 435
DB 456 SVYTLTATIERWHTITHAMQLDCKYQLRHAASVMWGMIFAFALPPIFGISSYMKVS 515
QY 436 LVLPSVILDLIQL 449
DB 516 ICLPMDIDSPLSQL 529

RESULT 10

FSHR_PIG STANDARD; PRT; 695 AA.
ID FSHR_PIG
AC P49059; 077514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
DE FSHR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96011644; PubMed=7590277;
RA Remy J.J., Labib-Mansais Y., Yertle W., Bozon V., Couture L., Pajot E., Grebert D., Salesse R.;
RT "The porcine follitropin receptor: cDNA cloning, functional expression and chromosomal localization of the gene.";
RL Gene 163:257-261(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F., la Barbera A.R.;
RT "Porcine follicle-stimulating hormone receptor.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity of this receptor is mediated by G proteins which activate adenylate cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors. FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DB EMBL; L31966; AAA86933.1; -.
DB EMBL; AF025377; AAC24981.1; -.
DB HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
KW phosphorylation; Repeat; Leucine-rich repeat.


```

FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT TRANSMEM 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT TRANSMEM 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CONFLICT 2 2
FT CONFLICT 13 13
FT CONFLICT 60 60
FT CONFLICT 166 166
FT CONFLICT 215 215
FT CONFLICT 247 247
FT CONFLICT 257 257
FT CONFLICT 334 334
FT CONFLICT 349 349
FT CONFLICT 352 352
FT CONFLICT 383 383
FT CONFLICT 407 407
FT CONFLICT 421 421
FT CONFLICT 427 427
FT CONFLICT 435 435
FT CONFLICT 483 483
FT CONFLICT 550 550
FT CONFLICT 586 586
FT CONFLICT 607 607
FT CONFLICT 691 691
SQ SEQUENCE 695 AA; 78172 MW; E9EBEDB29C79C450 CRC64;

Query Match 4.4%; Score 104.5; DB 1; Length 695;
Best Local Similarity 16.5%; Pred. No. 1.9;
Matches 91; Conservative 91; Mismatches 169; Indels 199; Gaps 21;

QY 18 LPNG-INGIKDARKVTGVTGSGDPAKSLTIRLIRCGYHVYIGSRNPKFASERFPVVDV 76
DB 61 IPKGAFFSGFDLEKI-----EISQNDVLEVEAN---VFSNLPKL----- 97
QY 77 THHEDALTKTNIIFFVAIHREHYTSLMDLRHLVG-----KILIDVSMN 119
DB 98 --HEIRIEKANNL-LYIDPDARQNLNRLRYLLISNTGVKHLPAVHKIQSLQKVLIDIQDN 154
QY 120 MRINQYPPESNAEYLASLFPDSLVKGFNVVSAWALQLPKADSKROVYICSNNIQARQCVI 179
DB 155 INIH-----TVERNSFVGLSFESMILWL---SKNGIREIHNCA----- 189
QY 180 ELARQINFIPIDLGLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPY 239
DB 190 -----FNGTQDLDELNLSNDNLEELLPNDFQGASGPVILDISRTRIHSLPSYLENLKCL 244
QY 240 ARNQSDPYKIP-----IEIVNKTU-----PIVAITLL----- 267
DB 245 RAKSTYVNLKKLPSLEKFTVTLMEASLTYPHSCCAFANWRROISDLHPICNKSIIRQEVDM 304

```

```

QY 268 -----SLVYIAGLLAAAYQLYGTYKRRF-----PPLETWLOCRKOLG--- 306
DB 305 TQARGQVSLAEDESSSLAKEFDYTWSEPDYDLQNEVVDVIGSEPDFTFNPCEDIMGHDI 364
QY 307 --LISFFPAMVHA-----YSLCLPMRRESERYLFILMAYQOVHANIENSWNE 351
DB 365 LRVLIWFISITAITGNITIVLVILTSQYKLTVP-----RFLMCNLAFPAD----- 408
QY 352 EEVWRIEMYISFGIMSLGLSLAVTISPSVSNALNPRE-----FSFIQSTLG--- 399
DB 409 -----LCIGIYLLLIASVDIHTKTYHNYAIDWQTGAGCDPAAGFTVFASELSVYT 459
QY 400 YVALLISTFH-----VLIIYGMKRAFEERYRFPYTPPNFV-LALVLP 439
DB 460 LTAITLERWHTITTHAMQLOCKVQLRHAASIMLVGWIFAFTVALPFIIGISSYMKVSIQCLP 519
QY 440 SIVILDLIQL 449
DB 520 MDIDSPISQL 529

```

```

RESULT 11
SOTB_ECO57
ID SOTB_ECO57 STANDARD; PRT; 396 AA.
AC P586529;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR Z2173 OR ECS2135.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
[1]
SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamosis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[2]
SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulon (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- SIMILARITY: Belongs to major facilitator superfamily. SotB
CC (TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```



```

CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE005352; AAG56236.1; -.
DR      EMBL; AP002557; BAB35558.1; -.
DR      PIR; G90895; G90895.
DR      PIR; H85721; H85721.
DR      HAMAP; MF_00517; -.
DR      InterPro; IPR007114; MFS.
DR      InterPro; IPR005828; Sub_transporter.
DR      Pfam; PF00083; sugar_tr; 1.
DR      PROSITE; PS50850; MFS; 1.
KW      Transport; Sugar transport; Transmembrane; Inner membrane;
KM      Complete proteome.
FT      TRANSMEM 15 35 POTENTIAL.
FT      TRANSMEM 50 70 POTENTIAL.
FT      TRANSMEM 81 101 POTENTIAL.
FT      TRANSMEM 103 123 POTENTIAL.
FT      TRANSMEM 136 156 POTENTIAL.
FT      TRANSMEM 170 190 POTENTIAL.
FT      TRANSMEM 209 229 POTENTIAL.
FT      TRANSMEM 246 266 POTENTIAL.
FT      TRANSMEM 275 295 POTENTIAL.
FT      TRANSMEM 299 319 POTENTIAL.
FT      TRANSMEM 333 353 POTENTIAL.
FT      TRANSMEM 364 384 POTENTIAL.
SQ      SEQUENCE 396 AA; 42568 MW; CB6A209B0BE6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY      31 VTIVGIVSGDFAKSLTIRLCGYHVIVIGSRNPKFASFPFPHVVDVTHHEDALTKNTIIF 90
DB      31 VPVGLL--SDIAQSTHMQTAQVGMILTIVAVYALMSLPFMLMTSGYERRKLLICLFVVF 88
QY      91 VAIHREHYTSLMDLRHLVVKILIDVSNMNRINQYPESNAEYTLASLPDLSLIVKGFNVVS 150
DB      89 IASHVLSFLS-WSFTVLVISRI-----GVAFAHAF-----WSITA 123
QY      151 AWALQLGP--KQASROYICSNNIQARQVIELAR-----QINFIPIDLGSL----- 195
DB      124 SLAIRMAPAGKRAQALSLIATGTALAMVLGPIGRIVGQYFGWRMTFPAIGALITLLC 183
QY      196 -----SSAREIENLPLRLFTLWRGPVVAISLAT-----FFLYSFVRDVIHPY 239
DB      184 LIKLPLLPSEHSGSLKSLPL---LFRPALMSIYLLTVVVVTAHYATASY---IEPF 235
QY      240 ARN---QQSDFYKIPIEIVNKLPIVAITLISLVYLAGLLAAAYQLYYGTYRRFPWLE 296
DB      236 VQNTAGFSANF-----ATALLLLIGAGIIGSVIFKLGNOYAS----- 274
QY      297 TWLQCRKQGLSLFPFPAWVAVSLCLPMSRSEKYLFLNMAVQGVHANIEIENSWNEEYVR 356
DB      275 -----ALVSTAIALLVCLALLPANSE-----IHUGVLSIF-----WG 309
QY      357 IEMYISFGIMSLGLSL-----LAVTSIPSVN-----ALNWRFFSFIQST 397
DB      310 IAMMIIGLGMQVKYVLLALPADATDVAMALFSGIFNIGIGAGALVGNQVSLHWS-----MSM 364
QY      398 LGYVALLISTFHVLYG-----WKRATFEFE 422
DB      365 IGYVG-TVPAFAPALIWIIIFRRMPVTLLEQ 394

RESULT 12
SOTB_ECOLI STANDARD; PRT; 396 AA.
AC P31122; P76883; P77353;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR B1528.

```

```

OS      Escherichia coli.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12;
RX      MEDLINE=97251357; PubMed=9097039;
RA      Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA      Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA      Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA      Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA      Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA      Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT      "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT      corresponding to the 28,0-40.1 min region on the linkage map.";
RL      DNA Res. 3:363-377(1996).
RN      [3]
RP      SEQUENCE OF 217-396 FROM N.A.
RX      MEDLINE=93186717; PubMed=8383113;
RA      Cohen S.P., Haechler H., Levy S.B.;
RT      "Genetic and functional analysis of the multiple antibiotic
RT      resistance (mar) locus in Escherichia coli.";
RL      J. Bacteriol. 175:1484-1492(1993).
RN      [4]
RP      CHARACTERIZATION.
RC      STRAIN=SB0;
RX      MEDLINE=99194728; PubMed=10094697;
RA      Bost S., Silva F., Belin D.;
RT      "Transcriptional activation of ydeA, which encodes a member of the
RT      major facilitator superfamily, interferes with arabinose accumulation
RT      and induction of the Escherichia coli arabinose PBAD promoter.";
RL      J. Bacteriol. 181:2185-2191(1999).
RN      [5]
RP      CHARACTERIZATION.
RC      STRAIN=JS219;
RX      MEDLINE=99369894; PubMed=10438792;
RA      Carole S., Pichoff S., Bouche J.-P.;
RT      "Escherichia coli gene ydeA encodes a major facilitator pump which
RT      exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside.";
RL      J. Bacteriol. 181:5123-5125(1999).
CC      -!- FUNCTION: Involved in the efflux of sugars. The physiological role
CC      may be the reduction of the intracellular concentration of toxic
CC      sugars or sugar metabolites. Transports L-arabinose and to a
CC      lesser extent IPTG. Seems to contribute to the control of the
CC      arabinose regulon.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC      (Probable).
CC      -!- SIMILARITY: Belongs to major facilitator superfamily. SotB
CC      (TC 2.A.1.2) family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE000250; AAC74601.1; -.
DR      EMBL; D90795; BAA15210.1; -.
DR      EMBL; D90796; BAA15218.1; -.
DR      EMBL; D90797; BAA15230.1; -.

```

DR EMBL; M96235; -, NOT_ANNOTATED_CDS.
DR PIR; C64907; C64907.
DR EcoGene; Egl1636; sotB.
DR HAMAP; MF_00517; -; 1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR pfam; PF00083; sugar_tr_1.
DR PROSITE; PS50850; MFS; 1.
DR TransPort; Sugar transporter; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42538 MW; CB6A34CA4EE6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

OY 31 VTGVIGSGDPAKSLTIRLCGHHVVISGRNPKFASEFPFHVDTTHEDALTKNIIIF 90
DB 31 VPGVLL--SDIAOSFHMOTQVGMILTYAAMVALMSLPFMTSVERKLLICLFYVF 88
OY 91 VAHREHYTSLMDLHLLVGKILIDVSNMNRINQYPESNAEYLASLPDLSLVKGNVVS 150
DB 89 IASHVLSFLS-WSFTVLVISRI-----GVAFHAIF-----WSITA 123
OY 151 AMALQLGP--KDSRQVYICSNNTQAPQVIELAR-----QNFIPIDLSL----- 195
DB 124 SLAIRMAPAGKRAQALSLIATGTALAMVLGLPLGRIVGQYFGWMTFFAIGALITLLC 183
OY 196 -----SSAREIENLPLRLFTLMRGPVVAISLAT-----FFLYSFVRDVIHPY 239
DB 184 LIKLPLLPBESHSGSLKSLPL--LERRPALMSIYLLTVVVVTAHTAYSX---IEPF 235
OY 240 ARN---QOSDFYKPIEIVNKTLPVIAITLISLVYLAGLAAAYQLYGTCKYRRFPWLE 296
DB 236 VQNIAGFSANF-----ATALLLLGAGIISVIFGLKGNQYAS----- 274
OY 297 TWLQCRKQLGLSFFFAMVHVAVSLCLPMRSEERYLFLNMAVQOVHANIENTSWNEEYVR 356
DB 275 -----ALVSTAIALLLVCLALLPAAENSE-----IHLGLSIF-----WG 309
OY 357 IEMYISFGIMSLGLSL-----LAVTSLPSVSN-----ALNWRSEFSFIQST 397
DB 310 IAMMIIGLGNQVKVLALAPDATDVAMALFSGIFNIGIGAGALVGNQVSLHWS-----MSM 364
OY 398 LGYVALLISTFHVLIYG-----WKRAFEFE 422
DB 365 IGYVG-AVPAFAALIWSTIFFRMPVTLSEQ 394

RESULT 13
NPT1_MOUSE
ID NPT1_MOUSE STANDARD; PRT; 465 AA.
AC Q61983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Renal sodium-dependent phosphate transporter protein 1 (Sodium/phosphate cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate transport protein 1) (Renal Na(+)-dependent phosphate cotransporter 1).
DE SLC17A1 OR NPT1.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95335846; PubMed=7611445;
RA Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T., Bourdeau J.E., Hughes M.R.;
RT "Cloning, genetic mapping, and expression analysis of a mouse renal sodium-dependent phosphate cotransporter.";
RL Am. J. Physiol. 268:F1038-F1045(1995).
CC -!- FUNCTION: Important for the resorption of phosphate by the kidney. May be involved in actively transporting phosphate into cells via Na(+)-cotransport in the renal brush border membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Kidney.
CC -!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X77241; CAAS4459.1; -.
DR PIR; S69915; S69915.
DR MGD; MGI:103209; SLC17A1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR004745; P1_cotranspt.
DR TIGRfam; TIGR00894; 2A0114euk; 1.
DR PROSITE; PS50850; MFS; 1.
KW TransPort; Symport; Sodium transporter; Transmembrane; Glycoprotein.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 47 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 51589 MW; C67EE25A2C291EEF CRC64;

Query Match 4.4%; Score 103; DB 1; Length 465;
Best Local Similarity 18.7%; Pred. No. 1.5; Indels 124; Gaps 16;
Matches 76; Conservative 64; Mismatches 142;

OY 110 GKILIDVSNMNRINQYPESNAEYLASLPDLSLVKGFNVVSAWALQLPKDA---SRQVY 166
DB 79 GLILSVFPGWVWVQAP--VGYISGIVPMKRITIGSSLFSLMSLLIPPAQVGALVI 135
OY 167 ICSNNIQARQVIELARQINFI-----PIDGLSSAREIENLPLRLFTLMRGPVV---- 218
DB 136 VCRVLQGIAGTGSTGHEIWVKNAPLERGLTS-----MTLSGFM--GPFIVLVLS 187
OY 219 -----AISLAFEEFLYSFVRDVHPYARNQOSDF----- 247
DB 188 GFICDLGWPWFYIFGIYGCVLSLSMFLLFPDDPKD--HPYMSSEKDYIISLMQAS 245
OY 248 ---YKPIEIVNKTLPVIAITLLS----- 268
DB 246 SGRQSLPIKAMLKSLPLMAIILNSFAFIWSNSLVVYTPPTISTVLHVNRENGLLSLP 305
OY 269 --LVYLAGLLAAAYQLYGTGKTRRFPFWLETWLQCRKQLGLISFFFAMVHVAVYSLCLPMR 326

Query Match 4.3%; Score 102; DB 1; Length 220;
Best Local Similarity 23.8%; Pred. No. 0.72;
Matches 43; Conservative 25; Mismatches 51; Indels 62; Gaps 7;

QY	267	LSLVYLAGLLAAAYQLYYGT-----KYRRFPFWLETW-----	298
Db	20	LMLLYTAGFVPAWTFYLGATGQLGADPVKTFEHLGLWALREFLLITLLVTPMRDLTGIT	79
QY	299	-LQCRKQGLISFFPAMVHVAYSCLPWRSEERYLFNNMAYQOVHANIENSWNEEEVWRI	357
Db	80	LLRYRRALGILLAFYYALMHFT-----TYMVLDOGINLSAITTDIVR-	120
QY	358	EMYISFGIMSLGILLSLAVTSIPSVSNALNW-----REFSFIQSTLGVALLISTFHYL	411
Db	121	RPFITIGMISLALVPLALTS-----NNWSIRKLGRRWSLHK-LVYIAIAGSAVHFL	172
QY	412	I 412	
Db	173	M 173	

Search completed: March 1, 2004, 23:31:14
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 23:34:47 ; Search time 63 Seconds
(without alignments)
1521.644 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMWSPKSLSETCLPN.....ALVLPSTVILDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2351	100.0	454	10	US-09-888-257A-10 Sequence 10, Appl
2	2351	100.0	454	14	US-10-165-044-8 Sequence 8, Appl
3	2351	100.0	454	15	US-10-239-607-29 Sequence 29, Appl
4	2351	100.0	454	15	US-10-295-027-574 Sequence 574, Appl
5	2351	100.0	454	15	US-10-392-190-2 Sequence 2, Appl
6	2348	99.9	454	15	US-10-104-047-3155 Sequence 3155, Ap
7	2294	97.6	490	9	US-09-802-520-1 Sequence 1, Appl
8	2290	97.4	490	15	US-10-239-607-14 Sequence 14, Appl
9	2036	86.6	419	15	US-10-239-607-32 Sequence 32, Appl
10	1272	54.1	488	14	US-10-182-951-6 Sequence 13, Appl
11	1269	54.0	488	9	US-09-747-835A-13 Sequence 36, Appl
12	1269	54.0	488	15	US-10-239-607-36 Sequence 38, Appl
13	1269	54.0	488	15	US-10-239-607-38 Sequence 21, Appl
14	1088	46.3	459	9	US-09-965-529-21 Sequence 21, Appl
15	1088	46.3	459	10	US-09-969-680A-21 Sequence 21, Appl

16	1087	46.2	459	14	US-10-165-044-10	Sequence 10, Appl
17	1073	45.6	461	15	US-10-239-607-34	Sequence 34, Appl
18	901	38.3	173	14	US-10-011-095-8	Sequence 8, Appl
19	901	38.3	173	14	US-10-010-667A-8	Sequence 8, Appl
20	846	36.0	375	10	US-09-895-298-162	Sequence 162, App
21	844	35.9	179	9	US-09-864-761-37654	Sequence 37654, A
22	768	32.7	237	9	US-09-747-835A-15	Sequence 15, Appl
23	736	31.3	141	9	US-09-963-896-1	Sequence 1, Appl
24	717	30.5	267	9	US-09-747-835A-50	Sequence 50, Appl
25	717	30.5	267	9	US-09-747-835A-51	Sequence 51, Appl
26	717	30.5	339	9	US-09-759-143-879	Sequence 879, App
27	717	30.5	339	9	US-09-780-669-879	Sequence 879, App
28	717	30.5	339	9	US-09-822-827-879	Sequence 879, App
29	717	30.5	339	9	US-09-802-520-11	Sequence 11, Appl
30	717	30.5	339	9	US-09-895-793-879	Sequence 879, App
31	717	30.5	339	9	US-09-895-814-879	Sequence 879, App
32	717	30.5	339	13	US-10-012-896-879	Sequence 879, App
33	717	30.5	339	14	US-10-011-095-2	Sequence 2, Appl
34	717	30.5	339	14	US-10-010-667A-2	Sequence 2, Appl
35	717	30.5	339	14	US-10-205-823-397	Sequence 397, App
36	717	30.5	339	14	US-10-144-678A-879	Sequence 879, App
37	717	30.5	339	14	US-10-294-025-879	Sequence 879, App
38	717	30.5	339	15	US-10-239-607-37	Sequence 37, Appl
39	717	30.5	339	15	US-10-295-027-714	Sequence 714, App
40	717	30.5	339	15	US-10-295-027-1347	Sequence 1347, Ap
41	717	30.5	375	14	US-10-165-044-2	Sequence 2, Appl
42	628	26.7	283	14	US-10-043-487-313	Sequence 313, App
43	628	26.7	283	14	US-10-205-823-123	Sequence 123, App
44	628	26.7	283	15	US-10-239-607-39	Sequence 39, Appl
45	433	18.4	232	10	US-09-895-298-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1

US-09-888-257A-10 Application US/09888257A

Sequence 10, Appl

Publication No. US20030060612A1

GENERAL INFORMATION:

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Polakis, Paul

APPLICANT: Smith, Victoria

APPLICANT: Wood, William I.

APPLICANT: Wu, Thomas D.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TITILE OF INVENTION: TREATMENT OF TUMOR

FILE REFERENCE: P5002R1

CURRENT APPLICATION NUMBER: US/09/888,257A

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/063,540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: US 60/089,653

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: US 60/099,792

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: US 60/103,678

PRIOR FILING DATE: 1998-10-08

PRIOR APPLICATION NUMBER: US 60/235,451

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: PCT/US99/12252

PRIOR FILING DATE: 1999-06-02

PRIOR APPLICATION NUMBER: PCT/US99/20111

PRIOR FILING DATE: 1999-09-01

PRIOR APPLICATION NUMBER: PCT/US00/04342

PRIOR FILING DATE: 2000-02-18

PRIOR APPLICATION NUMBER: PCT/US00/05841

PRIOR FILING DATE: 2000-03-02

PRIOR APPLICATION NUMBER: PCT/US00/08439

NOT PRIOR ART.

disclosed here

; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06666
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 10
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-888-257A-10

Query Match 100.0%; Score 2351; DB 10; Length 454;
Best Local Similarity 100.0%; Pred. No. 8e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKEASEFPFHVDTHEHDALTNTNIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
Db 61 RNPKEASEFPFHVDTHEHDALTNTNIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
QY 121 RINQYPESNAEYLALFPDLSLVKGFNVSAVALQGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLALFPDLSLVKGFNVSAVALQGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLSISAREIENLPLRLFTLMRGFVVAISLATFFFLYSFVRDVTHPYA 240
Db 181 LARQNFIPIDLSISAREIENLPLRLFTLMRGFVVAISLATFFFLYSFVRDVTHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTGRFRFPWLETWLQ 300
Db 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTGRFRFPWLETWLQ 300
QY 301 CRKQLGLSFFPAMVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENTSNWNEEVRRIEM 360
Db 301 CRKQLGLSFFPAMVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENTSNWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVALTSPSVSNALNWRRESFIQSTLGYVALLISTEHLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVALTSPSVSNALNWRRESFIQSTLGYVALLISTEHLIYGMKRAFE 420
QY 421 EHYRFTYTPNFVALVLPISIVILDLLQLCRYPD 454
Db 421 EHYRFTYTPNFVALVLPISIVILDLLQLCRYPD 454

RESULT 2
US-10-165-044-8
; Sequence 8, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Paris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 51158-20016-02
; CURRENT APPLICATION NUMBER: US/10/165,044
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01

; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-165-044-8

Query Match 100.0%; Score 2351; DB 14; Length 454;
Best Local Similarity 100.0%; Pred. No. 8e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKEASEFPFHVDTHEHDALTNTNIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
Db 61 RNPKEASEFPFHVDTHEHDALTNTNIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
QY 121 RINQYPESNAEYLALFPDLSLVKGFNVSAVALQGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLALFPDLSLVKGFNVSAVALQGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLSISAREIENLPLRLFTLMRGFVVAISLATFFFLYSFVRDVTHPYA 240
Db 181 LARQNFIPIDLSISAREIENLPLRLFTLMRGFVVAISLATFFFLYSFVRDVTHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTGRFRFPWLETWLQ 300
Db 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTGRFRFPWLETWLQ 300
QY 301 CRKQLGLSFFPAMVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENTSNWNEEVRRIEM 360
Db 301 CRKQLGLSFFPAMVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENTSNWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVALTSPSVSNALNWRRESFIQSTLGYVALLISTEHLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVALTSPSVSNALNWRRESFIQSTLGYVALLISTEHLIYGMKRAFE 420
QY 421 EHYRFTYTPNFVALVLPISIVILDLLQLCRYPD 454
Db 421 EHYRFTYTPNFVALVLPISIVILDLLQLCRYPD 454

RESULT 3
US-10-239-607-29
; Sequence 29, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchiglu, Fahri
; TITLE OF INVENTION: NO. US20030219761A1 Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 29
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapiens
US-10-239-607-29

Query Match 100.0%; Score 2351; DB 15; Length 454;
Best Local Similarity 100.0%; Pred. No. 8e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSFKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSFKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
DB 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
DB 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
QY 301 CRKQLGLSFFFAWVAVYSLCLPMSRERYLFNMAYQOVHANIENSWNEBEVWRIEMY 360
DB 301 CRKQLGLSFFFAWVAVYSLCLPMSRERYLFNMAYQOVHANIENSWNEBEVWRIEMY 360
QY 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPSTIVILDLLQLCRYPD 454
DB 421 EBYRFTYTPNPFVLAIVLPSTIVILDLLQLCRYPD 454

RESULT 4
US-10-295-027-574

Sequence 574, Application US/10295027
Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 574
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-574

Query Match 100.0%; Score 2351; DB 15; Length 454;
Best Local Similarity 100.0%; Pred. No. 8e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSFKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSFKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
DB 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
DB 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
QY 301 CRKQLGLSFFFAWVAVYSLCLPMSRERYLFNMAYQOVHANIENSWNEBEVWRIEMY 360
DB 301 CRKQLGLSFFFAWVAVYSLCLPMSRERYLFNMAYQOVHANIENSWNEBEVWRIEMY 360
QY 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPSTIVILDLLQLCRYPD 454
DB 421 EBYRFTYTPNPFVLAIVLPSTIVILDLLQLCRYPD 454

RESULT 5

US-10-392-190-2

Sequence 2, Application US/10392190
Publication No. US20040005598A1

GENERAL INFORMATION:

APPLICANT: Devaux, Brigitte
APPLICANT: Eberhard, David
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Hillan, Kenneth J.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Yansura, Daniel G.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: PUMPcN COMPOSITIONS AND USES THEREOF
FILE REFERENCE: P2994R1C1-US

; CURRENT APPLICATION NUMBER: US/10/392,190
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: US 60/235,451
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US01/30290
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-392-190-2

Query Match 100.0%; Score 2351; DB 15; Length 454;
Best Local Similarity 100.0%; Pred. No. 8e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
Db 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLMETWLQ 300
Db 241 RNQOSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLMETWLQ 300
QY 301 CRKQLGLSLFFFAVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
Db 301 CRKQLGLSLFFFAVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVALTSIPSVSNALNWRRESFIQSTLGYVALLISTHVLIIYWKRAFE 420
Db 361 ISFGIMSLGLSLAVALTSIPSVSNALNWRRESFIQSTLGYVALLISTHVLIIYWKRAFE 420
QY 421 EBYRFTYTPNFVLAFLPSIVILDLQLCRYPD 454
Db 421 EBYRFTYTPNFVLAFLPSIVILDLQLCRYPD 454

RESULT 6
US-10-104-047-3155
; Sequence 3155, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR APPLICATION NUMBER: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3155
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-104-047-3155

Query Match 99.9%; Score 2348; DB 15; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.6e-221;
Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
Db 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLMETWLQ 300
Db 241 RNQOSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPPLMETWLQ 300
QY 301 CRKQLGLSLFFFAVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
Db 301 CRKQLGLSLFFFAVHVAAYSLCLPMRRESERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVALTSIPSVSNALNWRRESFIQSTLGYVALLISTHVLIIYWKRAFE 420
Db 361 ISFGIMSLGLSLAVALTSIPSVSNALNWRRESFIQSTLGYVALLISTHVLIIYWKRAFE 420
QY 421 EBYRFTYTPNFVLAFLPSIVILDLQLCRYPD 454
Db 421 EBYRFTYTPNFVLAFLPSIVILDLQLCRYPD 454

RESULT 7
US-09-802-520-1
; Sequence 1, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Huel-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 7492448CD1
; US-09-802-520-1

Query Match 97.6%; Score 2294; DB 9; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.5e-216;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTVGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
Db 61 RNPKEASEFPFPHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIE 180

QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
DB 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
QY 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360
DB 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFTYTPNFVLAALVLPISIVIL 444
DB 421 EYYRFTYTPNFVLAALVLPISIVIL 444

RESULT 8
US-10-239-607-14
; Sequence 14, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchiogl'u, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-239-607-14

Query Match 97.4%; Score 2290; DB 15; Length 490;
Best Local Similarity 99.8%; Pred. No. 8.7e-216;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
DB 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
QY 61 RNPKFASEFPFHVVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
DB 61 RNPKFASEFPFHVVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQGPKASRQVYICSNNIQARQOYIE 180
DB 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQGPKASRQVYICSNNIQARQOYIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
DB 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
QY 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360
DB 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360

QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFTYTPNFVLAALVLPISIVIL 444
DB 421 EYYRFTYTPNFVLAALVLPISIVIL 444

RESULT 9
US-10-239-607-32
; Sequence 32, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchiogl'u, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-239-607-32

Query Match 86.6%; Score 2036; DB 15; Length 419;
Best Local Similarity 99.7%; Pred. No. 5.7e-191;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
DB 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
QY 61 RNPKFASEFPFHVVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
DB 61 RNPKFASEFPFHVVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVKGILIDVSNM 120
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQGPKASRQVYICSNNIQARQOYIE 180
DB 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAMALQGPKASRQVYICSNNIQARQOYIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFEFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
DB 241 RNOQSDFYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLO 300
QY 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360
DB 301 CRKQGLISFFPAMVHVAYSLCLPMRSEERYLFNMAYQOVHANINIENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQ 395
DB 361 ISFGIMSLGILLSLAVTSIPSVSNALNWRREFSFIQ 395

RESULT 10
US-10-182-951-6
; Sequence 6, Application US/10182951
; Publication No. US20030138895A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.


```

; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SHIH, Leo L.
; APPLICANT: YANG, Junming
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0033 PCT
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/181,856; 10/183,684; 60/185,141; 60/186,818; 60/188,345;
; PRIOR FILING DATE: 2000-02-11; 2000-02-17; 2000-02-25; 2000-03-03; 2000-03-09; 2000-0
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030138895A1 5629033CD1
US-10-182-951-6
```

```

Query Match 54.1%; Score 1272; DB 14; Length 488;
Best Local Similarity 54.5%; Pred. No. 6.4e-116;
Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDFAKSLTIRLRGCHVIGSRN 62
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 11 SLHLVSDSSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGFKVVGSRN 60
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 63 PKFASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSN--NM 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 61 PKRTARLPFAAQVTFQEEAVSSPEVIFAVFERHYSSLCSLSDQLAGKILVDVSNPTEQ 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQGPKDASROYVICSNNIOARQVTE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 121 EHLQHRSENAEYLAFLPCTCTVKAFTVISAATLQAGPRDGNRQVPCGDQPEAKRAVSE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 181 LARQNFIPIDLSLSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 241 RNOQSDFYKPIPIEIVNKTLPPIVAITLISLYLAGLLAAAYQLYGTYKRRPFWLETWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 QESQNKFFKLPVSVVNTTLPVAYVLLSLVYLLPGVLAALQLRGTGYQRFPDWLDMWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 301 CRKQGLLSFFPAMVHVAAYSLCLPMRSEERYLFLNMAYYQOVHANIEENSWNEEVRRIEMY 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 301 HRKQIGLLSFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRMEITY 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 361 LSLGVALLGTLISLAVTSLPSIANSLNWRSEFSFVQSSLGVALVSLHTLTLTYGWTSAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 421 EBYRYFTPPNFVALVLPISIVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 421 ESRYKFFLPPFTTLTLVPCVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

RESULT 11
US-09-747-835A-13
; Sequence 13, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```

; APPLICANT: Zhou, ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radjoe T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIK
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835A-13
```

```

Query Match 54.0%; Score 1269; DB 9; Length 488;
Best Local Similarity 54.3%; Pred. No. 1.3e-115;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDFAKSLTIRLRGCHVIGSRN 62
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 11 SLHLVSDSSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGFKVVGSRN 60
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 63 PKFASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSN--NM 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 61 PKRTARLPFAAQVTFQEEAVSSPEVIFAVFERHYSSLCSLSDQLAGKILVDVSNPTEQ 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQGPKDASROYVICSNNIOARQVTE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 121 EHLQHRSENAEYLAFLPCTCTVKAFTVISAATLQAGPRDGNRQVPCGDQPEAKRAVSE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 181 LARQNFIPIDLSLSAREIENTPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 241 RNOQSDFYKPIPIEIVNKTLPPIVAITLISLYLAGLLAAAYQLYGTYKRRPFWLETWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 QESQNKFFKLPVSVVNTTLPVAYVLLSLVYLLPGVLAALQLRGTGYQRFPDWLDMWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 301 CRKQGLLSFFPAMVHVAAYSLCLPMRSEERYLFLNMAYYQOVHANIEENSWNEEVRRIEMY 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 301 HRKQIGLLSFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRMEITY 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 361 LSLGVALLGTLISLAVTSLPSIANSLNWRSEFSFVQSSLGVALVSLHTLTLTYGWTSAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 421 EBYRYFTPPNFVALVLPISIVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 421 ESRYKFFLPPFTTLTLVPCVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

RESULT 12
US-10-239-607-36
; Sequence 36, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Saatchiglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1e1 Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-239-607-36
```

```
Query Match          54.0%; Score 1269; DB 15; Length 488;
Best Local Similarity 54.3%; Pred. No. 1.3e-115;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;
```

```
QY      3 SISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVGSRN 62
DB      11 SLHLVDSDSLAK--VPD-----EAPK--VGILGSGDFARSLATRLVSGFKVVGSRN 60

QY      63 PKFASFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSN--NM 120
DB      61 PKRTARLYPSAAQVTFOEAAVSSPEVIFVAVFREHYSSLCISLSDQLAGKILVDVSNPTQ 120

QY      121 RINQYPSNAEYLAFLPPDSLIYKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
DB      121 EHLQHRRESNAEYLAFLPPTCTVYKAFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180

QY      181 LARQINFIPIIDGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 MALANGFMPVDMGSLASAMEVEAMPRLLPAMKVPPTLALGLFVCFYAYNFVRDVLQPYV 240

QY      241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRRFPFMLETWLQ 300
DB      241 QESQNKPFKLPVSVVNTTLPVAYVLLSLVYLPGLAALQLRRGTXYQRFPMWLDHWLQ 300

QY      301 CRKQGLISFFEFAMVHVAYSCLPMRSEERYLFLNMAVOQVHANIENSWNEEVRRIEMY 360
DB      301 HRKQIGLISFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLMVEEVRMEIY 360

QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB      361 LSLGVLAGTLLSLAVTSLPSIANSINWREFSFVQSSLGVALVLSLHTLTFTGWTRAFE 420

QY      421 EBYRYFTYPNFVLAIVLPSIVIL 444
DB      421 ESRYKFTYLPPTFTLLVPCVIL 444
```

RESULT 13
US-10-239-607-38

```
; Sequence 38, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchiglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1e1 Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-239-607-38
```

```
Query Match          54.0%; Score 1269; DB 15; Length 488;
Best Local Similarity 54.3%; Pred. No. 1.3e-115;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;
```

```
QY      3 SISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVGSRN 62
DB      11 SLHLVDSDSLAK--VPD-----EAPK--VGILGSGDFARSLATRLVSGFKVVGSRN 60

QY      63 PKFASFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSN--NM 120
DB      61 PKRTARLYPSAAQVTFOEAAVSSPEVIFVAVFREHYSSLCISLSDQLAGKILVDVSNPTQ 120

QY      121 RINQYPSNAEYLAFLPPDSLIYKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
DB      121 EHLQHRRESNAEYLAFLPPTCTVYKAFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180

QY      181 LARQINFIPIIDGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 MALANGFMPVDMGSLASAMEVEAMPRLLPAMKVPPTLALGLFVCFYAYNFVRDVLQPYV 240

QY      241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRRFPFMLETWLQ 300
DB      241 QESQNKPFKLPVSVVNTTLPVAYVLLSLVYLPGLAALQLRRGTXYQRFPMWLDHWLQ 300

QY      301 CRKQGLISFFEFAMVHVAYSCLPMRSEERYLFLNMAVOQVHANIENSWNEEVRRIEMY 360
DB      301 HRKQIGLISFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLMVEEVRMEIY 360

QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB      361 LSLGVLAGTLLSLAVTSLPSIANSINWREFSFVQSSLGVALVLSLHTLTFTGWTRAFE 420

QY      421 EBYRYFTYPNFVLAIVLPSIVIL 444
DB      421 ESRYKFTYLPPTFTLLVPCVIL 444
```

RESULT 14
US-09-965-529-21

```
; Sequence 21, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preectl
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1859305CD1
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 19:57:15 ; Search time 83 seconds
(without alignments)
1725.847 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMWGSPKSLSETCLPN.....ALVLPSTIVLDLQLCRYPD 454

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2290	97.4	490	4	Q8IUE7	Q8IUE7 homo sapien
2	2284	97.2	490	4	Q8NFT2	Q8NFT2 homo sapien
3	2239.5	95.3	489	11	Q8BWB6	Q8BWB6 mus musculu
4	1318	56.1	488	11	Q8C5F0	Q8C5F0 mus musculu
5	1316	56.0	488	11	Q99P41	Q99P41 rattus norv
6	1316	56.0	488	11	Q8C159	Q8C159 mus musculu
7	1316	56.0	526	11	Q80ZF3	Q80ZF3 mus musculu
8	1306	55.6	514	11	Q924Z1	Q924Z1 mus musculu
9	1272	54.1	488	4	Q86SF6	Q86SF6 homo sapien
10	1272	54.1	498	4	Q7Z389	Q7Z389 homo sapien
11	1269	54.0	488	4	Q9NVB5	Q9NVB5 homo sapien
12	1252.5	53.3	487	4	Q8NEW6	Q8NEW6 homo sapien
13	1245.5	53.0	487	4	Q8TF03	Q8TF03 homo sapien
14	1074.5	45.7	456	4	Q8TDP3	Q8TDP3 homo sapien
15	1064.5	45.3	470	11	Q923B6	Q923B6 mus musculu
16	1057.5	45.0	470	11	Q91W31	Q91W31 mus musculu

17	1041.5	44.3	474	11	Q91ZE8	Q91ZE8 mus musculu
18	956.5	40.7	464	11	Q7TPB8	Q7TPB8 rattus norv
19	728	31.0	338	6	Q9GL50	Q9GL50 sus scrofa
20	715	30.4	339	11	Q924Z2	Q924Z2 mus musculu
21	707	30.1	339	11	Q9CWR7	Q9CWR7 mus musculu
22	704	29.9	339	11	Q924J9	Q924J9 mus musculu
23	628	26.7	283	4	Q9H5R1	Q9H5R1 homo sapien
24	541	23.0	283	4	Q8WMB0	Q8WMB0 homo sapien
25	324	13.8	143	4	Q9H7Y1	Q9H7Y1 homo sapien
26	171.5	7.3	208	16	Q8NLU6	Q8NLU6 corynebacte
27	170	7.2	208	16	Q8XQ52	Q8XQ52 ralstonia s
28	164.5	7.0	234	16	Q82AX0	Q82AX0 streptomyce
29	160	6.8	198	16	Q930K7	Q930K7 rhizobium m
30	159.5	6.8	211	16	Q8YK44	Q8YK44 anabaena sp
31	156.5	6.7	239	16	Q9KXR6	Q9KXR6 streptomyce
32	154.5	6.6	225	2	Q8GFG4	Q8GFG4 rhodococcus
33	151	6.4	200	16	Q98C62	Q98C62 rhizobium 1
34	149.5	6.4	218	2	Q8GFG3	Q8GFG3 rhodococcus
35	149.5	6.4	221	16	Q82Q14	Q82Q14 streptomyce
36	147.5	6.3	226	2	Q8GFK6	Q8GFK6 rhodococcus
37	147	6.3	213	17	Q29059	Q29059 archaeoglob
38	144.5	6.1	226	2	Q9AH05	Q9AH05 rhodococcus
39	143.5	6.1	223	17	Q8TYQ9	Q8TYQ9 methanopyru
40	142	6.0	212	17	Q29370	Q29370 archaeoglob
41	142	6.0	224	1	Q59661	Q59661 methanobact
42	137.5	5.8	199	2	Q9F418	Q9F418 mycobacteri
43	136.5	5.8	217	2	Q83VC1	Q83VC1 streptomyce
44	136.5	5.8	217	16	Q9JN78	Q9JN78 streptomyce
45	135.5	5.8	203	2	Q8GFG1	Q8GFG1 rhodococcus

ALIGNMENTS

RESULT 1

Q8IUE7
ID Q8IUE7 PRELIMINARY; PRT; 490 AA.
AC Q8IUE7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE STAMPl.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TRISU=Prostate;
RX MEDLINE=22229309; PubMed=12095985;
RA Korkmaz K.S., Ribi C.C., Korkmaz C.G., Loda M., Hager G.L.,
RA Saaticioglu F.;
RT "Molecular cloning and characterization of STAMPl, a highly prostate
RT specific six-trans-membrane protein that is overexpressed in prostate
RT cancer.";
RL J. Biol. Chem. 277:36689-36696(2002).
DR EMBL; AY008445; AAG32149.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 490 AA; 56027 MW; D1971A84880F7E51 CRC64;

Query Match	97.4%	Score 2290;	DB 4;	Length 490;
Best Local Similarity	99.8%	Pred. No. 4.1e-176;		
Matches 443;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MESISMWGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRICGHHVIGS	60	
DB	1	MESISMWGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRICGHHVIGS	60	
QY	61	RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM	120	
DB	61	RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM	120	

QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 180
Db 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 180
QY 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 300
Db 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 300
QY 301 CRKQLGLSLFFPAMVHVAYSLCLPMRSEERYLFLNMAYQOVHANIENSWNEEVEWRIEMY 360
Db 301 CRKQLGLSLFFPAMVHVAYSLCLPMRSEERYLFLNMAYQOVHANIENSWNEEVEWRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPISIVIL 444
Db 421 EBYRFTYTPNPFVLAIVLPISIVIL 444

RESULT 2

Q8NFT2 PRELIMINARY; PRT; 490 AA.
AC Q8NFT2; DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STEAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Porfka K.P., Helenius M.A., Visakorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
STEAP2, down-regulated in androgen-independent prostate cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF455138; AAN04080.1; -.
DR Genew; HGNC:17885; STEAP2.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
KW Transmembrane.
SQ SEQUENCE 490 AA; 55961 MW; F2E9C30CDACCEF81 CRC64;

Query Match 97.2%; Score 2284; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 1.2e-175;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNHVG 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNHVG 60
QY 61 RNPKEASEFFPHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
Db 61 RNPKEASEFFPHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 180
Db 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 180
QY 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

Db 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 300
Db 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 300
QY 301 CRKQLGLSLFFPAMVHVAYSLCLPMRSEERYLFLNMAYQOVHANIENSWNEEVEWRIEMY 360
Db 301 CRKQLGLSLFFPAMVHVAYSLCLPMRSEERYLFLNMAYQOVHANIENSWNEEVEWRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPISIVIL 444
Db 421 EBYRFTYTPNPFVLAIVLPISIVIL 444

RESULT 3

Q8BWB6 PRELIMINARY; PRT; 489 AA.
AC Q8BWB6; DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Weakly similar to tumor suppressor PHXDE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 95.3%; Score 2239.5; DB 11; Length 489;
Best Local Similarity 97.3%; Pred. No. 4.8e-172;
Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNHVG 60
Db 1 MESISMGSPKSL-ETFLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNHVG 59
QY 61 RNPKEASEFFPHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
Db 60 RNPKEASEFFPHVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 119
QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 180
Db 120 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYVICSNNIOARQVIE 179
QY 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 180 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 239
QY 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 300
Db 240 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGTGYKRRFPFWLETWLQ 299
QY 301 CRKQLGLSLFFPAMVHVAYSLCLPMRSEERYLFLNMAYQOVHANIENSWNEEVEWRIEMY 360

Db 300 CRKQGLSLFFFAVHVAVSLCLPMRKRERYLFLNMAYOOVHANIENTAMNEEYVRIEMY 359
QY 361 ISFGIMSLGLSLAVTSIPSVNALNWRPFSFIQSTIGVALLISTFHVLIGWKRAFE 420
Db 360 ISFGIMSLGLSLAVTSIPSVNALNWRPFSFIQSTIGVALLISTFHVLIGWKRAFA 419
QY 421 EBYRFFYTPNPFVLAIVLPSTIVIL 444
Db 420 EBYRFFYTPNPFVLAIVLPSTIVIL 443

RESULT 4

Q8C5F0 PRELIMINARY; PRT; 488 AA.
ID Q8C5F0
AC Q8C5F0;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE DUDULIN 2 homolog.
GN TSAP6 OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK078769; BAC37383.1; -
DR MGI; 1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 488 AA; 54780 MW; FDFCAF42AE503D11 CRC64;

Query Match 56.1%; Score 1318; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 8.4e-98;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MNGSPKSLSETCLPNGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGSRNPKF 65
Db 14 LVDSDSLAE--VP-----KEAPK--VGILSGSDFARSLATRLVSGFVVGSRNPKR 63
QY 66 ASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMRIN-- 123
Db 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSLADQLAGKILVDVSNPTEKEHL 123
QY 124 QYPESNAEYIASLFPDSLIVKGFNVSAWALQGPKASRQVYICSNNIQARQVIELAR 183
Db 124 QHRQSNABEYIASLFPACTVVKAFNVISAWALQAGPRDGNROVLICGQPEAKRTISEMAR 183
QY 184 QLNFTPIDGLSSAREIENTLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
Db 184 AMGFTPLDMGSLASAREVEAIFRLRLPSWKYPTLLALGLFVCFYTNFIRDLQPIRKD 243
QY 244 QSDFYKPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPFWLSTWLQCRK 303
Db 244 ENKFKYKMPLSVNTTIPCAVYVLLSLVYLPGLAALQLRGTGYQRFPPDLDHMLQHRK 303
QY 304 QUGLLSFFFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENTSNWEEVRIEMYISF 363
Db 304 QIGLLSFFFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENTSNWEEVRIEMYISL 363
QY 364 GIMSLGLSLAVTSIPSVNALNWRPFSFIQSTIGVALLISTFHVLIGWKRAFEERY 423
Db 364 GVALGLMGLSLAVTSIPSVIANSLNWKKEFSFVOSTLGFVALMLSTWHTLTLYGWTAFEEENH 423

QY 424 YRFYTPNPFVLAIVLPSTIVIL 444
Db 424 YRFYTPNPFVLAIVLPSTIVIL 444

RESULT 5

Q99P41 PRELIMINARY; PRT; 488 AA.
ID Q99P41
AC Q99P41;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RX MEDLINE=20424188; PubMed=10969787;
RA Steiner M.S., Zhang X., Wang Y., Lu Y.;
RT "Growth inhibition of prostate cancer by an adenovirus expressing a
RT novel tumor suppressor gene, pHyde."
RL Cancer Res. 60:4419-4425 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RA Lu Y., Rinaldy A.R., Steiner M.S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Rinaldy A.;
RT "A novel prostate cancer associated gene."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335281; AAK00361.1; -
DR EMBL; AF238865; AAL78207.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDCB CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MNGSPKSLSETCLPNGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGSRNPKF 65
Db 14 LVDSDSLAE--VP-----KEAPK--VGILSGSDFARSLATRLVSGFVVGSRNPKR 63
QY 66 ASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMRIN-- 123
Db 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSLADQLAGKILVDVSNPTEKERL 123
QY 124 QYPESNAEYIASLFPDSLIVKGFNVSAWALQGPKASRQVYICSNNIQARQVIELAR 183
Db 124 QHRQSNABEYIASLFPACTVVKAFNVISAWALQAGPRDGNROVLICGQPEAKRTISEMAR 183
QY 184 QLNFTPIDGLSSAREIENTLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
Db 184 AMGFTPLDMGSLASAREVEAIFRLRLPSWKYPTLLALGLSTQSYAYNFIRDLQPIRKD 243
QY 244 QSDFYKPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPFWLSTWLQCRK 303
Db 244 ENKFKYKMPLSVNTTIPCAVYVLLSLVYLPGLAALQLRGTGYQRFPPDLDHMLQHRK 303
QY 304 QUGLLSFFFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENTSNWEEVRIEMYISF 363
Db 304 QIGLLSFFFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENTSNWEEVRIEMYISL 363
QY 364 GIMSLGLSLAVTSIPSVNALNWRPFSFIQSTIGVALLISTFHVLIGWKRAFEERY 423
Db 364 GVALGLMGLSLAVTSIPSVIANSLNWKKEFSFVOSTLGFVALMLSTWHTLTLYGWTAFEEENH 423

QY 424 YRFYTPNFTLVLPISIVIL 444
DB 424 YKFYLPFTFTLLPCVITL 444

RESULT 6

Q8C159 PRELIMINARY; PRT; 488 AA.
AC Q8C159;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
DE Similar to RIKEN cDNA 1010001D01 gene.
GN TSAP6 OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037435; AAH37435.1; -.
DR MGD; MGI:1915678; Tgap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
SQ SEQUENCE 488 AA; 54749 MW; 9A08D99C90CF83F4 CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 14 LVDSDSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGFSVVGSRNPKR 63
QY 66 ASEPFPVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQHEAVSSPEVIFAVFREHYSSLSCLADQLAGKILVDVSNPTEKEHL 123
QY 124 QYPESNAEYLAFLPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIQARQVIELAR 183
DB 124 QHRQSNAEYLAFLPACTVVKAFNVISAMALQAGPRDGNRQVLICSDQPEAKRTISEMAR 183
QY 184 QLNFIPIIDGSLSSAREIENLPLRLFTLWRGPNVVAISLATPFPLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVPTLLALGLFVCFYTYNFIRDVLQPIRKD 243
QY 244 QSDFYKIPRIEIVNKTLPIVAITLSTVYLAGLLAAAYQLYGTGYRRFPPLWETWLOCRK 303
DB 244 ENKFYKMPLSVNTTLPVAVYVLLSVYLLPGVLAALQLRRGTYQRFPDWLDHWLOHRK 303
QY 304 QIGLLSFFFAVHVAYSCLPMRSEERYLFNMAVQOVHANINSEWNEEYWRMEIYLSL 363
DB 304 QIGLLSFFFAVHVAYSCLPMRSEERYLFNMAVQOVHANINSEWNEEYWRMEIYLSL 363
QY 364 GIMSLGLSLAVTSIPSVSNALNWEBSFIQSTLGVALLISTFHVLIYWKRAFEERY 423
DB 364 GVALGMLSLAVTSIPSVSNALNWEBSFIQSTLGVALLISTFHVLIYWKRAFEERY 423
QY 424 YRFYTPNFTLVLPISIVIL 444
DB 424 YKFYLPFTFTLLPCVITL 444

RESULT 7
Q80ZFB PRELIMINARY; PRT; 526 AA.
AC Q80ZFB;
DT 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-JUN-2003 (Tremblrel. 24, last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
DE Tsap6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ ACC331; TISSUE=Bone marrow;
RX MEDLINE=22506415; PubMed=12606722;
RA Passer B.J., Nancy-Portebois V., Amzallag N., Prieur S., Cans C.,
RA Roborel de Climens A., Fucci G., Bouvard V., Tynnder M., Susini L.,
RA Morchoisne S.P., Cribble V., Lespagnol A., Dausset J., Oren M.,
RA Amson R., Telerman A.;
RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell
RT cycle and interacts with Nix and the Myt1 kinase."
RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
DR EMBL; AY214462; AAC38239.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
SQ SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 526;
Best Local Similarity 56.2%; Pred. No. 1.3e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 52 LVDSDSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGFSVVGSRNPKR 101
QY 66 ASEPFPVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNMRIN-- 123
DB 102 TAGLFPSLAQVTFQHEAVSSPEVIFAVFREHYSSLSCLADQLAGKILVDVSNPTEKEHL 161
QY 124 QYPESNAEYLAFLPDSLIVKGFNVVSAMALQLGPKDASROYVICSNNIQARQVIELAR 183
DB 162 QHRQSNAEYLAFLPACTVVKAFNVISAMALQAGPRDGNRQVLICSDQPEAKRTISEMAR 221
QY 184 QLNFIPIIDGSLSSAREIENLPLRLFTLWRGPNVVAISLATPFPLYSFVRDVIHPYARNQ 243
DB 222 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVPTLLALGLFVCFYTYNFIRDVLQPIRKD 281
QY 244 QSDFYKIPRIEIVNKTLPIVAITLSTVYLAGLLAAAYQLYGTGYRRFPPLWETWLOCRK 303
DB 282 ENKFYKMPLSVNTTLPVAVYVLLSVYLLPGVLAALQLRRGTYQRFPDWLDHWLOHRK 341
QY 304 QIGLLSFFFAVHVAYSCLPMRSEERYLFNMAVQOVHANINSEWNEEYWRMEIYLSF 363
DB 342 QIGLLSFFFAVHVAYSCLPMRSEERYLFNMAVQOVHANINSEWNEEYWRMEIYLSL 401
QY 364 GIMSLGLSLAVTSIPSVSNALNWEBSFIQSTLGVALLISTFHVLIYWKRAFEERY 423
DB 402 GVALGMLSLAVTSIPSVSNALNWEBSFIQSTLGVALLISTFHVLIYWKRAFEERY 461
QY 424 YRFYTPNFTLVLPISIVIL 444
DB 462 YKFYLPFTFTLLPCVITL 482

RESULT 8
Q924Z1 PRELIMINARY; PRT; 514 AA.
AC Q924Z1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, last annotation update)
DE Dudulin 2.
GN TSAP6 OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Seriu V., Lambdin D., Lenoir C., Manivet P., Vaubourdolle M.,
RA Kellermann O., Loric S.;
RT "Molecular cloning and expression of mouse dudulin 2.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029586; AAK50539.1; -.
DR MGI; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 514 AA; 57268 MW; 339886C288AEC0E2 CRC64;

Query Match 55.6%; Score 1306; DB 11; Length 514;
Best Local Similarity 56.2%; Pred. No. 8.3e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVTGSGDPAKSLTIRLCGYHVVIGSRNPKF 65
Db 14 LVDSDSLAE--VP-----KEAPK--VGILGSGDFARSLATRLVSGSGFVVGSRNPKR 63
QY 66 ASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVGLIDVSNMRIN-- 123
Db 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYTSLCSLADQLAGKILVDVSNPTEKEHL 123
QY 124 QYEPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQVIELAR 183
Db 124 QHRQSNAEYLAFLPACTVVKAFNVISAWALQAVPRDGNRQVLIQNDKAKQRMEMAR 183
QY 184 QLNFIPIDLGSLASAREIENLPLRLFTLWRGPVVVAISLATEFFLYSFVRDVIHPYARNQ 243
Db 184 AMGFTPLDMGSLASAREVAIPLRLPSGKVPILLALGLFVCFYTYNFIIDVLCQPIRKD 243
QY 244 QSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPFWLQCRK 303
Db 244 ENKFKYKMPISLVNNTLPCVAYVLLSLVYLPGLAALQLRGTKYQRFPMIDHMLQHRK 303
QY 304 QLGILSFFPAMVHVAVSLCLPMRSEERYLFLNMAVQOVHANIEENSWNEEVRRIEMYISF 363
Db 304 QIGLISFFPAMVHALYSFCLPLRSHRYDLVNLAVKQVLANKSRLNVEEVRMEIYLSL 363
QY 364 GIMSGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 423
Db 364 GVALGMLSLAVTSLPSIANSINLWKEFSFVQSTLGFVALILSTMGTLTYGWTTRAFAEENH 423
QY 424 YRFYTPPNFVLAIVLPSIVIL 444
Db 424 YKFYLPPTFTLTLLPCVITL 444

RESULT 9

Q86SF6 PRELIMINARY; PRT; 488 AA.
ID Q86SF6;
AC Q86SF6;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to dudulin 2 (TSAP6).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22506415; PubMed=12606722;
RA Passer B.J., Nancy-Portebois V., Amzallag N., Prieur S., Cans C.,
RA Roborel de Climens A., Flucci G., Bouvard V., Tynnder M., Susini L.,

RA Morchoisne S.P., Crible V., Leespagnol A., Dausset J., Oren M.,
RA Amson R., Telexman A.;
RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell
RT cycle and interacts with Nix and the Myt1 kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
DR EMBL; BC042150; AA42150.1; -.
DR EMBL; AY214461; AAO38238.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 488 AA; 54600 MW; C89EB0D0430F9BFB CRC64;

Query Match 54.1%; Score 1272; DB 4; Length 488;
Best Local Similarity 54.5%; Pred. No. 4.3e-94;
Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTGSGDPAKSLTIRLCGYHVVIGSRN 62
Db 11 SLHVDSDSLAK--VPD-----EAPK--VGILGSGDFARSLATRLVSGSGFVVGSRN 60
QY 63 PKFASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120
Db 61 PKRTARLFPSSAAQVTFQEEAVSSPEVIFVAVFREHYTSLCSLADQLAGKILVDVSNPTEQ 120
QY 121 RINGYEPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQVIE 180
Db 121 EHLQHSNAEYLAFLPCTVVKAFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180
QY 181 LARQNFIPIDLGSLASAREIENLPLRLFTLWRGPVVVAISLATEFFLYSFVRDVIHPYA 240
Db 181 MALAMGMPVDMGSLASAMEVAMPRLLPAMKVPITLLALGLFVCFYANNFVRDVLQPYV 240
QY 241 RNQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPFWLQCRK 300
Db 241 QESQNKFFKLPVSVNNTLPCVAYVLLSLVYLPGLAALQLRGTKYQRFPMIDHMLQ 300
QY 301 CRKQGLSFFPAMVHVAVSLCLPMRSEERYLFLNMAVQOVHANIEENSWNEEVRRIEMY 360
Db 301 HRKQGLSFFPAMVHALYSFCLPLRSHRYDLVNLAVKQVLANKSRLNVEEVRMEIY 360
QY 361 ISFGMSGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db 361 LSLGVALGTLSLAVTSLPSIANSINLWREFSFIQSTLGFVALVSLTHTLTLYGWTTRAFAE 420
QY 421 ERYRFPNFPNFPVLAIVLPSIVIL 444
Db 421 ESRYKFPNFPNFPVLAIVLPSIVIL 444

RESULT 10

Q72389 PRELIMINARY; PRT; 498 AA.
ID Q72389;
AC Q72389;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686H07150.
GN DKFZP686H07150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endochel primary cell culture;
RA Bloecker H., Boecher M., Newes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538047; CAD97986.1; -.
KW Hypothetical protein.
SQ SEQUENCE 498 AA; 55593 MW; 69DD23D5A00C2D6B CRC64;

Query Match 54.1%; Score 1272; DB 4; Length 498;
Best Local Similarity 54.5%; Pred. No. 4.4e-94;
Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;

```
QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTGSGDFAKSLTIRLCGYHVVIGSRN 62
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 21 SLHLVSDSSSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGFKVVGSRN 70
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 63 PKFASEFFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 71 PKRTARLPSSAQVTFQEBAVSSPEVIFVAVREHYSSLCSLSDQLAGKITLVDSNPTEQ 130
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 RINQPESSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYICSNNIQARQVTE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 131 EHLQHRSSNAEYLAFLPCTCTVKAENVISAATLQAGPRDGNROYPICGDQPEAKRAVSE 190
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGFVVVAISLATFFFLYSFVRDVIHPYA 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 191 MALAMGFMPVDMGSLASAMEVEAMPRLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 250
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 241 RNQSDFFYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 251 QESQNKFFKLPVSVVNTTLPVAYVLLSLVYLLPGVLAALQLRGTXYQRFPDWLDHMLQ 310
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 301 CRKQLGLSFPFAMVHVAVYSLCLPMRSEERYLFLMAYQQVHANIEHNSWNEEVRRIEMV 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 311 HRKQIGLISFFCALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRMEIY 370
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 361 ISFGIMSLGLISLLAVTSIPSVSNALNWEFFSIQSTLGYVALLISTFHVLIYWKRAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 371 LSLGVLTALGTLISLLAVTSLPSIANSLNWREFFSVQSSLGVALVISTLHTLTYYGWTRAFE 430
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 421 EBYRFFYTPPNFVLAIVLPSIVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 431 ESRYKFFLPPFTTLLVPCVIL 454
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 11

Q9NVB5 PRELIMINARY; PRT; 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein FLJ10829 (Dudulin 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sertu V., Manivet P., Lenoir C., Eschwege P., Lambin D.,
RA Vaubourdolle M., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA91839.1; -.
DR EMBL; AY029585; AAK50538.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR03006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
KM Hypothetical protein.
SQ SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;

Query Match 54.0%; Score 1269; DB 4; Length 488;
Best Local Similarity 54.3%; Pred. No. 7.5e-94;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

```
QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTGSGDFAKSLTIRLCGYHVVIGSRN 62
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 11 SLHLVSDSSSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGFKVVGSRN 60
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 63 PKFASEFFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 61 PKRTARLPSSAQVTFQEBAVSSPEVIFVAVREHYSSLCSLSDQLAGKITLVDSNPTEQ 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 RINQPESSNAEYLAFLPDLSLVKGFNVSAWALQLGPKDASROYICSNNIQARQVTE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 121 EHLQHRSSNAEYLAFLPCTCTVKAENVISAATLQAGPRDGNROYPICGDQPEAKRAVSE 180
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGFVVVAISLATFFFLYSFVRDVIHPYA 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 241 RNQSDFFYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 QESQNKFFKLPVSVVNTTLPVAYVLLSLVYLLPGVLAALQLRGTXYQRFPDWLDHMLQ 300
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 301 CRKQLGLSFPFAMVHVAVYSLCLPMRSEERYLFLMAYQQVHANIEHNSWNEEVRRIEMV 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 301 HRKQIGLISFFCALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRMEIY 360
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 361 ISFGIMSLGLISLLAVTSIPSVSNALNWEFFSIQSTLGYVALLISTFHVLIYWKRAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 361 LSLGVLTALGTLISLLAVTSLPSIANSLNWREFFSVQSSLGVALVISTLHTLTYYGWTRAFE 420
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 421 EBYRFFYTPPNFVLAIVLPSIVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 421 ESRYKFFLPPFTTLLVPCVIL 444
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 12

Q8NEW6 PRELIMINARY; PRT; 487 AA.
AC Q8NEW6;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Phye.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshti B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hPhye, that
RT inhibits prostate cancer cell growth."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082673; AAM45136.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR03006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54501 MW; E0EF95E855C81EDF CRC64;

Query Match 53.3%; Score 1252.5; DB 4; Length 487;
Best Local Similarity 54.1%; Pred. No. 1.6e-92;
Matches 240; Conservative 83; Mismatches 108; Indels 13; Gaps 5;

```
QY 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVTGSGDFAKSLTIRLCGYHVVIGSRN 62
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 11 SLHLVSDSSSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGFKVVGSRN 60
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 63 PKFASEFFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```



```
Db      61 PKRTARLFPSAQAQVTFQOEAVSSPEVIFVAVFREHYSSLCISLSDQLAGKILVDVSNPTEQ 120
Qy      121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
Db      121 EHLQHRRESNAEYIASLFPCTVVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180
Qy      181 LARQINFIPIIDGLSSASAREIENTPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPRLLPAMKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
Qy      241 RNOQSDFYKPIEIVNKTLPPIVATILSLVYLAGLLAAAYQLYGTKYRRFPPLWLETWLQ 300
Db      241 QESQNKFFKLPVSVVNTTLPVAVYLLSLVYLLPGVLAALQLRGTQYQRFPDWLDHMLQ 300
Qy      301 CRKQGLLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Db      301 HRKQIGLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Qy      361 ISFGIMSLGLISLLAVTSIPSVSNALNWRFFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
Db      360 LSLGVALLGTLSLLAVTSLPSIANSLNWRFFSFVQSSLGVALVSLTLHTLTYGWTRAFE 419
Qy      421 EBYRFFYTPPNFVALVLPISIVIL 444
Db      420 ESRYKFYLPPTFTLTLLVPCVIL 443
```

RESULT 13

```
Q8TF03 PRELIMINARY; PRT; 487 AA.
AC Q8TF03;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Allay J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54402 MW; AFP16053590E6F68 CRC64;
```

Query Match 53.0%; Score 1245.5; DB 4; Length 487;
Best Local Similarity 53.8%; Pred. No. 5.8e-92;
Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

```
Qy      3 SISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLIRCGYHVVIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGGFKVAVGSRN 60
Qy      63 PKFASFPFHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGLKILIDVSN--NM 120
Db      61 PKRTARLFPSAQAQVTFQOEAVSSPEVIFVAVFREHYSSLCISLSDQLAGKILVDVSNPTEQ 120
Qy      121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
Db      121 EHLQHRRESNAEYIASLFPCTVVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180
Qy      181 LARQINFIPIIDGLSSASAREIENTPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPRLLPAMKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
```

```
Qy      241 RNOQSDFYKPIEIVNKTLPPIVATILSLVYLAGLLAAAYQLYGTKYRRFPPLWLETWLQ 300
Db      241 QESQNKFFKLPVSVVNTTLPVAVYLLSLVYLLPGVLAALQLRGTQYQRFPDWLDHMLQ 300
Qy      301 CRKQGLLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Db      301 HRKQIGLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Qy      361 ISFGIMSLGLISLLAVTSIPSVSNALNWRFFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
Db      360 LSLGVALLGTLSLLAVTSLPSIANSLNWRFFSFVQSSLGVALVSLTLHTLTYGWTRAFE 419
Qy      421 EBYRFFYTPPNFVALVLPISIVIL 444
Db      420 ESRYKFYLPPTFTLTLLVPCVIL 443
```

RESULT 14

```
Q8TDP3 PRELIMINARY; PRT; 456 AA.
AC Q8TDP3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE pHyde II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Allay J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF262322; AAM08128.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 456 AA; 50430 MW; C5F7C7008D55251E CRC64;
```

Query Match 45.7%; Score 1074.5; DB 4; Length 456;
Best Local Similarity 52.9%; Pred. No. 3.2e-78;
Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

```
Qy      3 SISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLIRCGYHVVIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGGFKVAVGSRN 60
Qy      63 PKFASFPFHVVDVTHEDALTKNIIFVAIHREHYTSLWDLRHLVGLKILIDVSN--NM 120
Db      61 PKRTARLFPSAQAQVTFQOEAVSSPEVIFVAVFREHYSSLCISLSDQLAGKILVDVSNPTEQ 120
Qy      121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
Db      121 EHLQHRRESNAEYIASLFPCTVVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180
Qy      181 LARQINFIPIIDGLSSASAREIENTPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPRLLPAMKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
Qy      241 RNOQSDFYKPIEIVNKTLPPIVATILSLVYLAGLLAAAYQLYGTKYRRFPPLWLETWLQ 300
Db      241 QESQNKFFKLPVSVVNTTLPVAVYLLSLVYLLPGVLAALQLRGTQYQRFPDWLDHMLQ 300
Qy      301 CRKQGLLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Db      301 HRKQIGLSFFFCALHALYSFCLPLBRAHRYDLVNLAVKQVLANKSHLW-VEEVMRMEIY 359
Qy      361 ISFGIMSLGLISLLAVTSIPSVSNALNWRFFSFIO 395
Db      360 LSLGVALLGTLSLLAVTSLPSIANSLNWRFFSFVQ 394
```

RESULT 15

Q923B6 PRELIMINARY; PRT; 470 AA.
AC Q923B6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to hypothetical protein FLJ23153 (Tnfa-induced adipose-related protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RM [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; BC006651; AAH06651.1; -.
DR EMBL; AK040760; BAC30696.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR004455; NADPoxred.F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00024; HEMOPEXIN; 1.
SQ SEQUENCE 470 AA; 52994 MW; 6823E7682AC78B19 CRC64;

Query Match 45.3%; Score 1064.5; DB 11; Length 470;
Best Local Similarity 48.5%; Pred. No. 2.le-77;
Matches 200; Conservative 80; Mismatches 131; Indels 1; Gaps 1;

QY 33 VGVIGSGDEFAKSLTIRLCGYHVVIGSRNPKFSEFFPHVVDVTHHEDALTKTNIIFVA 92
DB 22 VCIFGTGDFGSLGKMLQCGYSIVFGSRNPQ-VSSLLPRGAEVLSYSEASKSDIITLA 80
QY 93 IHREHYTSLWDLRHLVGVKILIDVSNRRINQYPESNAEYLASLPDLSLIVKGFNVSAW 152
DB 81 MREHYDSLTELVDYLGKVLVDVSNRRINQYPESNAEYLASLPDLSLIVKGFNVSAW 140
QY 153 ALQGPXDASROYVICSNNIOARQOVIELARQINFIPIIDLSISAREIENLPLRLFTLW 212
DB 141 ALQSGTLDASROYVCGNDSKAKQRYMDIARTLGLTPLDQSLMAASRIENYPLQLFPMW 200
QY 213 RGPVVVAISLATFEFLYSFVRDVIHPYARNQOSDFYKIPLEIVNKTLPVAVITLSIVYL 272
DB 201 RFPFYLSSVLCVFFFYCAIREVITYPVNGKTDATYRLAISIPNRVFPITALLIALVYL 260
QY 273 AGLLAAAYQLYGTXYRRFPFWLETLQCRKQLGLLSFFFAWVAVASLCLPMRRSERYL 332
DB 261 PGILAILQLYRGTKYRRFPNWLDMWLCKRQLGLVALGFAFLHVTYTLVPIRYVWR 320
QY 333 FLNMAYQOVHANIEENSWNEEVRWRIEMYSFGIMSLGLLSLAVTSIPSVSNALNWRFS 392
DB 321 LKNATITQALTNDSPFITSYAMINDSYALGILGFLFLIGITSLSPSVSNMVMWRFR 380
QY 393 FIQSTLGVVALLISTFHVLIYGVKRAFESEYRFTYTPNFVLAIVLPSIVIL 444
DB 381 FVQSKLGYLTIVLCTAHTLVYGGKRFSLPSILRWSLPSAYITLAVIPCAYLV 432

Search completed: March 1, 2004, 23:34:42
Job time : 89 secs